

Db	241	gpqitltlhtdltvpcliciqvwplpdsrvtnlopfredprahgnlwgqarlrlltllqswl	3000
Qy	301	LDAPCSLPAEALCALCRAAGSGPPCPPLVPLPSWENTVD-----VNSSE	3433
Db	301	ldapcslpaeealcvtrapsgqpcplvplswentvdkvlefplkhgnplcvgnasse	3606
Qy	344	KLLOECIMADSLAPLKKDVLLEFRGQDNRSICALEPSCGSLPSKASRARLGEYL	4033
Db	361	kldqecimadslaplkadvlllefrgqdnrsicalpysgtslpskasttraarlgelyl	4200
Qy	404	LDLDSGOCLOLMDDDLALMACFMDKYLHRRNALWLIACLLFAAALSLILLLKKDKAKG	4633
Db	421	ldldsgocclqlmddldlgaicwcpmdkylhkrvalwliacllfaaalsllllkkdkakg	4800
Qy	464	WLRLKQVRSAGAAARCAALLLTYSADDSGFERLVGALASALCOLPLRAVADLMSRREL	5233
Db	481	wlrlkqvrsagaaraaalltlysaddsgferlvgalasalcqlplrvavdlwstrrels	5400
Qy	524	AQGPAPWHAQRQTLOEGGVVLLFSGANALCSEWLQDSVSPGAHGPHDAFRASLSC	5833
Db	541	aqgpapwhaqrqtllqegvgvvllfsgaavalcsewldqdvsgpghdpdfraslsc	6000
Qy	584	VLPPLQCRAPGSYVGAFCEDLTPDAVPALFRVVPFTLPSOLPDLGLAQPRAPRSG	6433
Db	601	vlpplqcrapgsyvgafcdllpdavpalftrvvpftlpsqlpdlglalqpraprsq	6600
Qy	644	RLQERAEQVSRALQPALDSTYHPGTFAPAGRGVPGAGPGAGDGT	688
Db	661	rlqeraeqvsralqpaldstyhpgtcfapagrvpgpagaagdggt	705
RESULT 4			
AAB87606			
ID	AAB87606	standard; Protein; 705 AA.	
XX	AC		
XX	AAB87606;		
XX	DT	15-MAY-2001 (first entry)	
XX	DE	Human PRO20040.	
XX	KM	Human; PRO protein; mapping.	
XX	OS	Homo sapiens.	
XX	XX	WO200116318-A2.	
XX	PD	08-MAR-2001.	
XX	FE	24-AUG-2000; 2000MO-US23328.	
XX	PR	01-SEP-1999; 99MO-US20111.	
XX	PR	15-SEP-1999; 99MO-US21090.	
XX	PR	07-DEC-1999; 99US-0169495.	
XX	PR	09-DEC-1999; 99US-0170762.	
XX	PR	11-JAN-2000; 2000US-0175481.	
XX	PR	18-FEB-2000; 2000MO-US04341.	
XX	PR	18-FEB-2000; 2000MO-US04342.	
XX	PR	22-FEB-2000; 2000MO-US04414.	
XX	PR	01-MAR-2000; 2000MO-US05601.	
XX	PR	03-MAR-2000; 2000US-0187202.	
XX	PR	25-APR-2000; 2000US-0199397.	
XX	PR	22-MAY-2000; 2000MO-US14042.	
XX	PR	05-JUN-2000; 2000US-0209832.	
XX	PA	(GETH) GENENTECH INC.	
XX	PI	Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ,	
XX	PI	Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI.	
XX	PR	WPI; 2001-183260/18.	
XX	PR	N-PSDB; AAF92138.	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 17:01:01 ; Search time 43.62 Seconds

(without alignments)
610.707 Million cell updates/sec

Title: US-09-608-918-10
Perfect score: 3692
Sequence: 1 MPVFWFLSLALGRSPVLS.....TPAPGRGVGRAGGAGDGT 688

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128.5	3.5	1711	1	PTPO_RAT
2	126	3.4	3707	1	PGBM_MOUSE
3	124.5	3.4	761	1	POOF_KLEPN
4	113.5	3.1	587	1	NDC2_RAT
5	109.5	3.0	428	1	NER3_HUMAN
6	108.5	2.9	753	1	UR06_HSVB
7	107	2.9	825	1	IL4R_HUMAN
8	106.5	2.9	1709	1	SN_HUMAN
9	105.5	2.9	741	1	TLE4_RAT
10	105	2.8	766	1	TLE4_HUMAN
11	102	2.8	1208	1	RCO4_HUMAN
12	101.5	2.7	418	1	NER3_RAT
13	100	2.7	1011	1	MAK6_HUMAN
14	99	2.7	505	1	CXAA_MOUSE
15	99	2.7	922	1	W70T_MOUSE
16	98.5	2.7	890	1	TYO3_HUMAN
17	98	2.7	5430	1	ACE7_HUMAN
18	97.5	2.6	676	1	UR06_HSV1
19	97.5	2.6	850	1	NRG2_HUMAN
20	97.5	2.6	897	1	CYRB_HUMAN
21	97	2.6	993	1	AFSR_STRCO
22	96.5	2.6	992	1	SNX1_HUMAN
23	95	2.6	528	1	VRP3_GIALA
24	95	2.6	577	1	TLE4_MOUSE
25	94.5	2.6	209	1	CXX1_HUMAN
26	94	2.5	392	1	GAG_BLYAU
27	94	2.5	2205	1	POLN_RUBYT
28	94	2.5	4393	1	PGBM_HUMAN
29	93.5	2.5	459	1	CAH9_HUMAN
30	93.5	2.5	775	1	POOF_PSEAE
31	93.5	2.5	880	1	TYO3_RAT
32	93.5	2.5	1442	1	PTC1_CAEEL
33	92.5	2.5	379	1	YP85_CAEEL

34	92.5	2.5	537	1	TCAL_MOUSE	P13597 mus musculus
35	92.5	2.5	867	1	SSPO_BOVIN	P68167 bos taurus
36	92.5	2.5	947	1	M3KE_HUMAN	O99538 homo sapien
37	92.5	2.5	1473	1	NAL1_HUMAN	O9C000 homo sapien
38	92	2.5	443	1	W70T_RAT	O35828 rattus norv
39	92	2.5	492	1	CPSL_PIG	P15540 sus scrofa
40	92	2.5	553	1	MIS_RAT	P49000 rattus norv
41	92	2.5	672	1	NOO3_PARDE	P29915 paracoccus
42	92	2.5	843	1	Y096_CAEEL	P41846 caenorhabd1
43	92	2.5	1705	1	PTPO_MOUSE	P70289 mus musculu
44	90.5	2.5	368	1	GALT_HUMAN	O60755 homo sapien
45	90.5	2.5	508	1	EPOR_HUMAN	P19235 homo sapien

ALIGNMENTS

RESULT 1
PTPO_RAT STANDARD; PRT; 1711 AA.
AC 064612;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE Osteosticellular protein tyrosine phosphatase precursor (EC 3.1.3.48)
DE (OST-PTP).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Osteosarcoma;
RX MEDLINE=95074080; PubMed=7527035;
RA Mauro L.J., Olmsted E.A., Skrobacz B.M., Mourey R.J., Davis A.R.,
RA Dixon J.E.;
RT "Identification of a hormonally regulated protein tyrosine
RT phosphatase associated with bone and testicular differentiation.";
RL J. Biol. Chem. 269:30659-30667(1994).
CC -1- FUNCTION: MAY FUNCTION IN SIGNALING PATHWAYS DURING BONE
CC REMODELING, AS WELL AS SERVE A BROADER ROLE IN CELL INTERACTIONS
CC ASSOCIATED WITH DIFFERENTIATION IN BONE AND TESTIS. OPTIMAL PH FOR
CC PHOSPHATASE ACTIVITY IS 5.6.
CC -1- ASSOCIATED WITH DIFFERENTIATION IN BONE AND TESTIS.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: A PRESUMED ALTERNATE TRANSCRIPT OF 4.8-5.0
CC KILOBASES, WHICH MAY LACK PTP DOMAINS, IS PRESENT IN PROLIFERATING
CC OSTEOBLASTS, BUT NOT DETECTABLE AT OTHER STAGES.
CC -1- TISSUE SPECIFICITY: BONE AND TESTIS. IN THE LATTER, RESTRICTED TO
CC THE BASAL PORTION OF THE SEMINIFEROUS TUBULE.
CC -1- DEVELOPMENTAL STAGE: UP-REGULATED IN DIFFERENTIATING CULTURES OF
CC PRIMARY OSTEOBLASTS AND DOWN-REGULATED IN LATE STAGE MINERALIZING
CC CULTURES. IN TESTIS, EXPRESSION IS HIGHEST BETWEEN STAGES I AND
CC VII WHEN MATURING SPERMATIDS REMAIN BURIED WITHIN THE SERTOLI
CC EPITHELIUM.
CC -1- INDUCTION: BY PARATHYROID HORMONE AND CYCLIC AMP ANALOGS.
CC -1- PTM: THE CYTOPLASMIC DOMAIN CONTAINS POTENTIAL PHOSPHORYLATION
CC SITES.
CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC -1- SIMILARITY: CONTAINS 10 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: L36884; AAA63911.1; -
CC DR HSSP: P18052; 1YFO.

```

DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003962; FNIII_repeat.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; TYR_prot_phphatase.
DR Pfam; PF00041; fo3; 7.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00700; PRTYPHPHATASE.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00194; PTPC; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 2.
KW Hydrolase; Transmembrane; Repeat; Signal; Glycoprotein.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 1711 OSTEOCYTICULAR PROTEIN TYROSINE
FT DOMAIN 18 1074 PHOSPHATASE.
FT TRANSMEM 1075 1095 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1096 1711 CYTOSOLIC (POTENTIAL).
FT DOMAIN 32 124 FIBRONECTIN TYPE-III 1.
FT DOMAIN 125 215 FIBRONECTIN TYPE-III 2.
FT DOMAIN 216 303 FIBRONECTIN TYPE-III 3.
FT DOMAIN 304 392 FIBRONECTIN TYPE-III 4.
FT DOMAIN 393 470 FIBRONECTIN TYPE-III 5.
FT DOMAIN 471 562 FIBRONECTIN TYPE-III 6.
FT DOMAIN 563 652 FIBRONECTIN TYPE-III 7.
FT DOMAIN 653 741 FIBRONECTIN TYPE-III 8.
FT DOMAIN 742 830 FIBRONECTIN TYPE-III 9.
FT DOMAIN 831 921 FIBRONECTIN TYPE-III 10.
FT DOMAIN 1150 1418 PROTEIN-TYROSINE PHOSPHATASE 1.
FT DOMAIN 1419 1711 PROTEIN-TYROSINE PHOSPHATASE 2.
FT ACT_SITE 1350 1350 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 431 431 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 570 570 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 663 663 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 851 851 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 882 882 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 970 970 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 982 982 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1711 AA; 187292 MW; 0F04D2D1A47A18A0 CRC64;

Query Match 3.5%; Score 128.5; DB 1; Length 1711;
Best Local Similarity 20.0%; Pred. No. 0.067;
Matches 159; Conservative 74; Mismatches 268; Indels 295; Gaps 39;

QY 5 WFLLSALG-RSPVLSLERLVGPQDATHGSPGLSRLWSDILCLPGDIVAPGPVLAP 63
DB 322 WASRWAGLARGDGYVL---KLSGPMESTSTLGPPECN-----AVFPGP-LPP 364
QY 64 THLQTLVLRCKQETDCDLCLRVAVHLVGHGHEEPEDEKFGGAADSGVEE---PRNA 119
DB 365 GHYITQL-----KVLGAPYDAWVEGSGTWLAE 392
QY 120 SLQAGV-----LSFOAYPTARCVLLEVQVPAALVQFGSQVGS--VYDCEFAALGS 169
DB 393 ALPREVPGARLWLDGLEASKQGRALLYSDDAPGSLGNISVPSGATHVIFCGLVPGAHY 452
QY 170 EVRWISYTOPRYKELNHTQQLP--ALPWLNVSDAGDNVHLVLNVSEQHFGLSLYWNQV 227
DB 171 EVRWISYTOPRYKELNHTQQLP--ALPWLNVSDAGDNVHLVLNVSEQHFGLSLYWNQV 227

```

```

Db 453 RVDIASSTGDISQISGYSTPLPPQSLVSRSSPSD-----LTIANGPA 497
QY 228 QCP---PKRWHKNT-----GQIITLNHTDLP--CLCIQVWP-----LEPDSV 268
Db 498 PGQLEGYKVTWHQDGSQSPGDLVDLQPDTLSTLTKSLVPGCYTVSAWAWAGNLDSDSQ 557
QY 269 RTNCP-----FREDPRAHONLW-----QAARLRLTLQ----- 297
Db 558 KIHSTRPAPTNTLSLGFHAQPAALKASWYHPGGRDAPFLRLYRLPTLSEKYLPRE 617
QY 298 ---SWL-LDAPCSLPAEALCW-----RAPGDPCQPLVPPLSWENVVDVNSSEK 344
Db 618 AQNFSMAQLTAGCEQVQLSTLWGSERSSANATGTP--PSAPTL-----VNTSDAP 569
QY 345 LQQLQCLWADSLGPKLDDVLL-----ETR-----GP-QDNRSICALEPS-----GCTSLP 389
Db 670 TQLQ-VSWAHVPGGRSRYQVTLYQESTRTATSIMGPKEDGTSLGLTPTGKYKVEISWA 728
QY 390 SKASTRAARLGEY-----LLQDLQSGCQLQMLDDDLGALWACPMCKYIHKRWALVWL 441
Db 729 GPLYTAANVSAWYPLIPNELLVSMQAGS-----W-----PSGPLQGACACHAQSADAGHLSEW 789
QY 442 ACLLEFAAALSLLLLKKDHAKGWLRLKODVBSGAARALLLYSADDSGFERLVGAL 501
Db 764 A-----W-----PSGPLQGACACHAQSADAGHLSEW 789
QY 502 ASALCQLPLRVAVDLWSRELSSAQGVAVFHAQRRTLOEG-GVVVLLFSPGAVALLCSEW 560
Db 790 -----FLKGOELFMLRDLTPGHTISVRCRAGPLQASTHLVLSVEPGPV----- 836
QY 561 LODGVSGPAGHPDAFRASLSCVLPD-----FLOGRAGPS-----YVACFDRLL 606
Db 837 -EDVLCHEP-----EATYALALNWTMPAGDVCLVVRVVERLVPGGGTHFVQVNTSGDALL 889
QY 607 HPDAPALFRTVPVFTLPSQLPDFLQALQOPRAPSGRQLQERAEQVSRALQALDSYFHP 666
Db 890 -LPNLPTTYSRLSTVL-----GRNSRVSRAVSLVCSAEAW-HP 929
QY 667 PGTPAP-----GRGVG 677
Db 930 PELAEPPQVVELGTGMG 945

RESULT 2
PGBM_MOUSE
ID PGBM_MOUSE STANDARD; PRT; 3707 AA.
AC Q05793;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (Perlecan) (PLC).
GN HSPG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=92078153; PubMed=1744087;
RA Noonan D.M., Fullen A., Valente P., Cai S., Horigan E., Sasaki M.,
RA Yamada Y., Hassell J.R.;
RT "The complete sequence of perlecan, a basement membrane heparan
RT sulfate proteoglycan, reveals extensive similarity with laminin A
RT chain, low density lipoprotein-receptor, and the neural cell adhesion
RT molecule."
RL J. Biol. Chem. 266:22939-22947(1991).
RN [2]
RP SEQUENCE OF 940-1601 AND 1870-2600 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=89034110; PubMed=2972708;
RA Noonan D.M., Horigan E.A., Ledbetter S.R., Vogeli G., Sasaki M.,
RA Yamada Y., Hassell J.R.;

```



```
FT DISULFID 1211 1234 BY SIMILARITY.
FT DISULFID 1237 1246 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1275 1287 BY SIMILARITY.
FT DISULFID 1277 1293 BY SIMILARITY.
FT DISULFID 1295 1304 BY SIMILARITY.
FT DISULFID 1307 1322 BY SIMILARITY.
FT DISULFID 1363 1372 BY SIMILARITY.
FT DISULFID 1365 1379 BY SIMILARITY.
FT DISULFID 1582 1591 BY SIMILARITY.
FT DISULFID 1594 1610 BY SIMILARITY.
FT DISULFID 1613 1628 BY SIMILARITY.
FT DISULFID 1615 1638 BY SIMILARITY.
FT DISULFID 1641 1650 BY SIMILARITY.
FT DISULFID 1653 1668 BY SIMILARITY.
FT DISULFID 1792 1839 BY SIMILARITY.
FT DISULFID 1886 1932 BY SIMILARITY.
FT DISULFID 1976 2021 BY SIMILARITY.
FT DISULFID 2073 2118 BY SIMILARITY.
FT DISULFID 2170 2215 BY SIMILARITY.
FT DISULFID 2268 2313 BY SIMILARITY.
FT DISULFID 2365 2413 BY SIMILARITY.
FT DISULFID 2456 2506 BY SIMILARITY.
FT DISULFID 2554 2599 BY SIMILARITY.
FT DISULFID 2641 2686 BY SIMILARITY.
FT DISULFID 2631 2876 BY SIMILARITY.

Query Match 3.4%; Score 126; DB 1; Length 3707;
Best Local Similarity 19.6%; Pred. No. 0.29;
Matches 97; Conservative 41; Mismatches 111; Indels 246; Gaps 25;

Qy 4 PWFLSLALGRSPVLSRLVGPQDAT-----HCSP-----GLSC----- 39
Db 272 PQFLPSVPGPS-----ACGPEASCHSGHCIPRDYLCDDQEDCRDGSDELGCASP 323
Qy 40 -----RLWDSILCLPGDI-----VPAGCPVLAPTHLOTEL 71
Db 324 PCEPNEFACENGHCALKWRCD-----GDFCEDRTDEANCSVKQPGVCGPTHFCVST 378
Qy 72 LR-----COKETCDCLRLVAVLHVHWEPEDEEKFGEAAGDSGVVEEPNLSIOAQ 124
Db 379 NRCIPASPHCDSEDC-----PDRSEFGCMPPQVVTTPQOQ-SIOA- 418
Qy 125 VVLSFOAYPTARCVLLEQVPAALVQFGOSGVSVYDCEFAALGSEVRWISVTPRYEKE 184
Db 419 -----SRGQ-----TVFTCTVATGVPTPIINW-----R 441
Qy 195 LNHTQQLPALPWLNVASGDNVHLNVSEQHFG-----LSLYW 224
Db 442 LN-WGHIPAPRPVMTSEGGRTLIIRDVKADQAGAYTCEAMNSRGWFGIPDGVLV- 499
Qy 225 NOVQGPAPRWKHLNLTGQIITLHNTDLYPCLIQWPLEPDSVRTNICP-----FREDPR 280
Db 500 --PORGCPDGHFYLE-----DSASCLPCFCFGY-----TNVQSSLRFRDQIR 541
Qy 281 A---HONLQWQARLRLLTLQSWLLDAPCSLPAEALCWAPGDCPCQPLVPPLSWENVTV 337
Db 542 LSFDPNDFKGVNVTM-----PSQGVGPPPLSSTQLQI 573
Qy 338 D-----VNSSEKLOQECILWA-----DSLQ-----PL-KD 361
Db 574 DPALQEQFLVDLSRRELVDHAFWALPKQPLGNKVDYSYGGFLRVKRYELARGMLEPVQXP 633
Qy 362 DVLL-----LETRG 370
Db 634 DVILWAGYRLHSRG 648

RESULT 3
POOF_KLEPN
ID POOF_KLEPN STANDARD; PRT; 761 AA.
AC P27508;
DT 01-AUG-1992 (Rel. 23, Created)
```

```
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Coenzyme PQQ synthesis protein F (EC 3.4.99.-).
GN POOF.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
[1]
RP SEQUENCE FROM N.A.
RN STRAIN=NCTC 418;
RX MEDLINE=92212293; PubMed=1313537;
RA Meulenbergh J.J.M., Sellink E., Riegman N.H., Postma P.W.;
RT "Nucleotide sequence and structure of the Klebsiella pneumoniae poq
operon.";
RL Mol. Gen. Genet. 232:284-294(1992).
CC -I- FUNCTION: REQUIRED FOR COENZYME PYROLO-QUINOLINE-QUINONE (POQ)
CC BIOSYNTHESIS. IT IS THOUGHT THAT THIS PROTEIN IS A PROTEASE THAT
CC CLEAVES PEPTIDES BOND IN A SMALL PEPTIDE (GENE POQA) WHICH WOULD
CC PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH ARE NECESSARY
CC FOR THE SYNTHESIS OF POQ.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
CC INSULINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58778; CAA41584.1; -
DR PIR; S20458; S20458.
DR MEROPS; M16.006; -
DR InterPro; IPR001431; Peptidase_M16.
DR Pfam; PF00675; Peptidase_M16; 1.
DR PROSITE; PS00143; INSULINASE; 1.
KW POQ; Hydrolase; Metalloprotease; Zinc.
FT METAL 49 49 ZINC (BY SIMILARITY).
FT ACT_SITE 52 52 BY SIMILARITY.
FT METAL 53 53 ZINC (BY SIMILARITY).
FT METAL 130 130 ZINC (BY SIMILARITY).
SQ SEQUENCE 761 AA; 83616 MW; 187182DBC8E839B5 CRC64;

Query Match 3.4%; Score 124.5; DB 1; Length 761;
Best Local Similarity 24.5%; Pred. No. 0.049;
Matches 120; Conservative 47; Mismatches 171; Indels 151; Gaps 27;

Qy 282 HONLQWQARLRLLTLQSWLLDAPCSLPAEALCWAPGG---DPCQPLVPPLSWENVTV 338
Db 182 HRHYVARMQL-----W-LQGQSLALGELAARFAAGLAAGAPPPAPPLRGFTA- 234
Qy 339 VNSSEKLOQ-----QECLEWADSLGPKDDVLLLETRGPDNRSICALPEPGCTSLPSKAST 394
Db 235 -----LQLAVSSOPALWRCPLIALSDNVLLR----- 261
Qy 395 RAARLGEYLLQD-----LQSGQCQLQWDDDLGALWACPMKYIHKRWLVWLACLLFAAL 450
Db 262 -----EFLDEAPGSLMAGLRORRLRAGDVALNW-----LYODRHLGWLALVFAADR 307
Qy 451 SLILLKKKHAKGWLRLKQDV-----RSGAAA-----RGRALLYSADDS 492
Db 308 PEEVDROITH---WLQALQOTTPEOQQHYTQLSRRRQALSPLDLQRLQRAFGAPGAPPA 364
Qy 493 GFERLVGAL-----ASALCQ-----LPLRVAVDWLSRRELSSAQGPVAMFHA 533
Db 365 GFADFCALQVAPSVSLACQTVSPGEPVATGGFSLPL-----AVALCSEWLQDGV- 420
Qy 534 QRQRTL-----QEGGVVVLFFSPG-----AVALCSEWLQDGV- 567
Db 421 QAAGDLVAKCEKAAPLLHLPSFGDPPRLRLRPFFYCSPOAQEGLARGEQLRPLLAALRH 480
```

```

QY 568 PGARPHDAFRAS--LSCVLPDFLOGRAGSVYACFDRLHMDA--VPALEFRTVPEVTL 623
DB 481 AGGHEEMHFLFDSMQLTTLQPE--PCRREAILQALITRLALFVASLTPSP--PSAIRHL 537
QY 624 PSQULDFGALGOQPR---APRSGRLOERAEQVSRALQPALDSFHHPCGPAPG---RGVG 677
DB 538 MAQLPERLGTSHQKGMALALAGSAEDHOMVAKOLS-LITAFVNP--MPAPACRRGVE 595
QY 678 PGAGPGAGD 686
DB 596 RUYPG-GD 603

RESULT 4
NDCC2_RAT STANDARD: PRT: 587 AA.
AC P70545:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Intestinal sodium/dicarboxylate cotransporter (Na(+)/dicarboxylate cotransporter).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MITAR;
RA MEDLINE=97107437; PubMed=8950177;
RT "Cloning of the cDNA for a rat intestinal Na+/dicarboxylate cotransporter reveals partial sequence homology with a rat intestinal mucin."
RT cotransporter reveals partial sequence homology with a rat intestinal mucin."
RL Biochim. Biophys. Acta 1309:58-62(1996).
CC -1- FUNCTION: COTRANSPORT OF SODIUM AND DICARBOXYLATES SUCH AS SUCCINATE AND CITRATE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE NADC/P/PHO87 FAMILY OF TRANSPORTERS. NADC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as there is no content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U51153; AAB97095.1;
DR InterPro; IPR001898; Na_sulph_symp.
DR Pfam; PF00939; Na_sulph_symp; 3.
DR PROSITE; PS01271; NA_SULFATE; 1.
RW Transport; Transmembrane; Sodium transport; Symport.
FT TRANSMEM 13 33 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 326 346 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 407 427 POTENTIAL.
FT TRANSMEM 445 465 POTENTIAL.
FT TRANSMEM 477 497 POTENTIAL.
FT TRANSMEM 506 526 POTENTIAL.
FT TRANSMEM 535 555 POTENTIAL.
SO SEQUENCE 587 AA; 63531 MW; F9B74F921BDC8712 CRC64;

Query Match 3.1%; Score 113.5; DB 1; Length 587;
Best Local Similarity 21.3%; Pred. No. 0.26;
Matches 91; Conservative 47; Mismatches 166; Indels 123; Gaps 20;
QY 312 ALCWRAPGGDP-----CQPL-VPLSWENVTVVNSSEKIQLOECIMADSLGRLKDDVL 364

```

```

DB 2 ATCPALMANRFTYILVLCPLFLPLIYOTKAYCAYSITLALLCTEALPLAVTVL 61
QY 365 -----LEFTRGPONRSICALPEPSCGTSIPSKASTRAARLGEYLIQDLOSQCQLQMDDDL 420
DB 62 FPVILFLMGIMD-----ASBGLHRLSRTPI-----YCLSGADGHCCTLP-- 105
QY 421 GALTACPMQDYIHK--RNALVWLACLLPAAALSL-----TLLKKDHA-KGWLRLIKQ 470
DB 106 -AQTCHPSPASYSRASAALLLLGFMVLVAFLSMISMTATTAMVPIGHAEVLQASKK 164
QY 471 DVRSQ-----
DB 165 DVEGGNNPFELOECPQKQEVTKLDNGQPVASBPSTQKTOGHHRFSQSLCTCYSA 224
QY 490 DDSGFERYLGALASALCOLPLRVAVDLSRRELASQGVAMFHAORROLOEGGVLLF 549
DB 225 SIGIDNLTGTTPNLVLOQOVN-----SLFPRKQCGELCEVVFRLPHDHLAAQ---LAM 278
QY 550 SPGAVALC---SEWL-----QDVSQPGAGHPHDAFRASLVCVLPDFLO 590
DB 279 ATGPLPCQQLPEELMLNGRGRTKACGEPGHQDPVQAAG---HEFCREDLFTVLV 335
QY 591 ----GRAGSVG---ACF-----DRLHPDAVPLFRTVPTLPQSLPDLGLAQQPRA 639
DB 336 VLWFTREPGFPGGDTVFANEGKQSMPSDGTVAIFSLVWFIIPSKIP--GLMEDPK 392
QY 640 PRSGRLQ 646
DB 393 P--GKTK 397

RESULT 5
NER3_HUMAN STANDARD: PRT: 428 AA.
AC Q9UQA9; Q9NOEL;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Stialidase 3 (EC 3.2.1.18) (Membrane stialidase) (ganglioside stialidase) (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99335353; PubMed=10405317;
RA Wada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;
RT "Cloning, expression, and chromosomal mapping of a human ganglioside stialidase."
RL Biochem. Biophys. Res. Commun. 261:21-27(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX PubMed=10861246;
RA Monti E., Bassi M.T., Papini N., Riboni M., Manzoni M., Venerando B., Croci G., Preti A., Ballabio A., Tettamanzi G., Borsani G.;
RT "Identification and expression of NEU3, a novel human stialidase associated to the plasma membrane."
RL Biochem. J. 349:343-351(2000).
CC -1- FUNCTION: Plays a role in modulating the ganglioside content of the lipid bilayer at the level of membrane-bound stialyl glycoconjugates.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-, alpha-(2->8)-glycosidic linkages of terminal sialic residues in oligosaccharides, glycoproteins, glycolipids, colominic acid and synthetic substrates.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- TISSUE SPECIFICITY: Highly expressed in skeletal muscle, testis, adrenal gland and thymus, followed by pancreas, liver, heart and

```

CC thymus. Weakly expressed in kidney, placenta, brain and lung.
 CC -!- MISCELLANEOUS: Optimum pH is 3.8.
 CC -!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.
 CC -!- SIMILARITY: CONTAINS 3 BNR REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB008185; BAA82611.1; -
 DR EMBL; Y18563; CAB96131.1; ALT_INIT.
 DR MIM; 604617; -
 DR InterPro; IPR002860; BNR.
 DR Pfam; PF02012; BNR; 3.
 KW Hydrolase; Glycosidase; Membrane; Repeat.
 FT REPEAT 129 140 BNR 1.
 FT REPEAT 203 214 BNR 2.
 FT REPEAT 254 265 BNR 3.
 FT SITE 24 27 FRIP MOTIF.
 FT ACT_SITE 25 25 BY SIMILARITY.
 FT ACT_SITE 45 45 POTENTIAL.
 FT ACT_SITE 50 50 POTENTIAL.
 FT ACT_SITE 87 87 POTENTIAL.
 FT ACT_SITE 225 225 POTENTIAL.
 FT ACT_SITE 245 245 POTENTIAL.
 FT ACT_SITE 340 340 BY SIMILARITY.
 FT ACT_SITE 370 370 POTENTIAL.
 FT ACT_SITE 387 387 POTENTIAL.
 SQ SEQUENCE 428 AA: 48252 MW: 35D1DD9359A78C98 CRC64;

Query Match 3.0%; Score 109.5; DB 1; Length 428;
 Best Local Similarity 21.1%; Pred. No. 0.35;
 Matches 107; Conservative 44; Mismatches 158; Indels 199; Gaps 26;
 QY 170 EVRIWSTQPP--RYEKELNHTQOLPALPWL-----NVSADGNVHLVNLVSE 214
 DB 3 EVTSCFNSPLRQEDDRGITHIRIPALLYPHTFLAFKRSRDRDEALHLVLR--R 60
 QY 215 EQHFGLSLYWNOVQPPKPRWKNLTGPOIITLHNDLVPCLCQVWPLPDSVRTNICP 274
 DB 61 GLRIGOLVQW---GELKPLMEATLPG-----HRTMNP-----PWQXSGCVFLEFFIC 106
 QY 275 FREDPRAHQNL---WQAAKRLRLTLQSWLLDAPCSLPAAALCWRAPGDPCQPLVPPLS 331
 DB 107 VRGHVTERQQIVSGRNAARLCFYQ-----DAGC-----S 137
 QY 332 WENV---TVDVNSSEKLOQCLWAD-SLGP-----LKDDVILLETGPGQDNBSLCALEP 382
 DB 138 WSEVRDLTEEVIGSELKH-----WATFVAGPGHGHIQSGRLVIP-----AYTYIIP 184
 QY 383 SG--CTSPLPSKASTRAARLGEYLLQDLQSGQCQLQWDDDLGALWACPMKYIHKRW--AL 438
 DB 185 SWFFCFQLPCKTSPHS-----LMYISDDLGVTH-----HGRLIRPM 221
 QY 439 VWLAC-----LFFAALS-----LILLKKDHAKGWLRL----- 467
 DB 222 VTVEVEAEBVTGRAGHPVLVYCSARTPNRCRAEALSTDHGEFGFORLALSRQCEPPHCQCG 281
 QY 468 -----LKQDVRSGAARAGRAALLIYSADDSGF 494
 DB 282 SVVSRPLPIPHRCODSSSKDAPTIOQSPGSSRLREEEAGTPSE---SWLLYSHTPSRK 338
 QY 495 ERLVGALASALCOLPLRVAVDLSWRSLSAQGPVAFHQAQRROTLOEGGVVLLFSPGAV 554
 DB 339 QRV--DLGIYLNQTPLEAA--CWSRWIILHCGPCG---SDLAALKEEGLFGCLFECGTK 391
 QY 555 ALCSE-----WLQDVGSGPG 569

Db 392 QECEQIAFRLEFTHREILSHLOQDCTSPG 419
 RESULT 6
 UL06_HSVB STANDARD; PRT; 753 AA.
 AC P28944; 1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Virion gene 56 protein.
 GN 56.
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 RN NCBI_TaxID=31520;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92295566; PubMed=1318606;
 RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
 RL "The DNA sequence of equine herpesvirus-1";
 RL Virology 189:304-316(1992).
 CC -!- FUNCTION: PRESUMED VIRION PROTEIN; POSSIBLE ROLE IN DNA
 CC PACKAGING.
 CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL6,
 CC EHV-1 56, EBV BBF1, HCMV UL104, AND VZV 54.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M86664; AAB02491.1; -
 DR PIR; B36801; WZBEE8.
 DR InterPro; IPR002660; Herpes_UL6.
 DR Pfam; PF01763; Herpes_UL6; 1.
 DR PRODOM; PD003210; Herpes_UL6; 1.
 SQ SEQUENCE 753 AA: 83992 MW: C5E118F78BBED203 CRC64;
 Query Match 2.9%; Score 108.5; DB 1; Length 753;
 Best Local Similarity 22.1%; Pred. No. 0.88;
 Matches 100; Conservative 46; Mismatches 164; Indels 143; Gaps 22;
 QY 295 TLQSWLLDAPCSLPAAALCWRA--PGGD-----PCOPLVP-----P 329
 DB 156 TLQSWLLDFVRSITA---CFASSEPDCGTASFAYIDWIACLGLIPLQLKRAPGATVHP 211
 QY 330 LSWENVTVDVNSSEKLOQCLWADSLGLKDDVLLLTETGPGQDNBSL----- 377
 DB 212 KLMKRLPTDVPS-----LESCV-----DERDLIAGKLYVANSLLR 245
 QY 378 ---CALEPSGCT-SLPSKASTRAARLGEYLLQDL-----OSGQCLQCLWD---DDLQ 421
 DB 246 EGLEAVVELACTASVIMDYDRVNIFFHYTRREVAIDSTTGKRGECCLVWQPIWKDGS 305
 QY 422 ALWACPMCKY-----IHKRWALVWLACLLFFAALSILLKLDHAKGWLRLKQDVRSG 475
 DB 306 VLFDSPLOKRGCEVCNCHALREHAKLCOLLNTVPVKILVGRKKDEAQG-----PG 355
 QY 476 AAARGRAALLIYSADDSGFERLVGALASALCOLPLRVAVDLSWRSLSAQGPVAFHQAQR 535
 DB 356 WASKAVDKLM-----GEHEELHSSSAASRL---VKLIWNKSMRHIGDITETVRSYLN 406
 QY 536 RQTLQEGGVVLLFSPGAVALCSWLQDVGSGPGAHGP-HDAFRASLSCVLPDFLOGRAP 594
 DB 407 TSTNLLSGAQVDYTSLPG-----FGOSGKTKQGNMPPVQEAFTSVINGNMLEG--- 456
 QY 595 GSVYGACFDRLHHPDPAVPAFTVPVFTLPQSLPDFLQALQOPRAPRSRGLQRAEQVSR 654

DB 457 --YVN-----NLKFTIEDLRNGS-----GLLDQLR-DESEITHLREQLLR 495
OY 655 ALQPALDSTFHPPGPA--PGRCVGPAGAG 685
DB 496 VSOAADGSTOPGASSALPGSGAKSGAG-GLG 527

RESULT 7
IL4R_HUMAN STANDARD: PRT; 825 AA.
AC P24394;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Interleukin-4 receptor alpha chain precursor (IL-4R-alpha) (CD124 antigen).
GN IL4R OR IL4RA OR 582J2.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Peripheral blood;
RX MEDLINE-90171849; PubMed-23079314;
RA Idzerda R.L., March C.J., Mosley B., Lyman S.D., Bos T.V., Gimpel S.D., Din W.S., Grabstein K.H., Widmer M.B., Park L.S., Cosman D., Beckmann M.P.;
RT "Human interleukin 4 receptor confers biological responsiveness and defines a novel receptor superfamily.";
RL J. Exp. Med. 171:861-873(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91120547; PubMed-2278997;
RA Gallizzi J.-P., Zuber C.E., Harada N., Gorman D.M., Djossou O., Kastlein R., Bancheau J., Howard M., Miyajima A.;
RT "Molecular cloning of a cDNA encoding the human interleukin 4 receptor.";
RL Int. Immunol. 2:669-675(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-99425270; PubMed-10493829;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).
RN [4]
RP VARIANTS VAL-75; ALA-400; ARG-431; LEU-436 AND PRO-786.
RX MEDLINE-97224413; PubMed-9070874;
RA Delchmann K., Bardutzky J., Forster J., Heinzmann A., Kuehr J.;
RT "Common polymorphisms in the coding part of the IL4-receptor gene.";
RL Biochem. Biophys. Res. Commun. 231:696-697(1997).
RN [5]
RP VARIANT ATOPIC ARG-576.
RX MEDLINE-98041803; PubMed-9392697;
RA Hershey G.K.K., Friedrich M.F., Esswein L.A., Thomas M.L., Chaitila T.A.;
RT "The association of atopy with a gain-of-function mutation in the alpha subunit of the interleukin-4 receptor.";
RL New Engl. J. Med. 337:1720-1725(1997).
RN [6]
RP VARIANT ATOPIC ASTHMA VAL-75.
RX MEDLINE-96282087; PubMed-9620765;
RA Mitsuysu H., Izuhara K., Mao X.-Q., Gao P.S., Arinobu Y., Enomoto T., Kawai M., Sasaki S., Date Y., Hamasaki N., Shirakawa T., Hopkin J.M.;
RT "The 50val variant of IL4R alpha upregulates Ige synthesis and associates with atopic asthma.";
RL Nat. Genet. 19:119-120(1998).
RN [7]
RP VARIANT ATOPIC ASTHMA VAL-75.

RX MEDLINE-99322293; PubMed-10390422;
RA Noguchi E., Shimasaki M., Arinami T., Takeda K., Yokouchi Y., Kobayashi K., Imoto N., Nakahara S., Matsui A., Hamaguchi H.;
RT "No association between atopy/asthma and the Ile50val polymorphism of IL-4 receptor.";
RL Am. J. Respir. Crit. Care Med. 160:342-345(1999).
RN [8]
RP VARIANTS PRO-503 AND ARG-576.
RX MEDLINE-99250314; PubMed-10233717;
RA Kruse S., Japha T., Tedner M., Sparholt S.H., Forster J., Kuehr J., Delchmann K.A.;
RT "The polymorphisms S503P and Q576R in the interleukin-4 receptor alpha gene are associated with atopy and influence the signal transduction.";
RL Immunology 96:365-371(1999).
RN [9]
RP VARIANT ALA-752.
RX MEDLINE-20143377; PubMed-10677312;
RA Ober C., Leavitt S.A., Tsalenko A., Howard T.D., Hokl D.M., Daniel R., Newman D.L., Wu X., Parry R., Lester L.A., Solway J., Blumenthal M., King R.A., Xu J., Meyers D.A., Blecker E.R., Cox N.J.;
RT "Variation in the interleukin 4-receptor alpha gene confers susceptibility to asthma and atopy in ethnically diverse populations.";
RL Am. J. Hum. Genet. 66:517-526(2000).
RN [10]
RP VARIANT ATOPIC ARG-576.
RX MEDLINE-20269830; PubMed-10809862;
RA Oiso N., Fukui K., Ishii M.;
RT "Interleukin 4 receptor alpha chain polymorphism Gln551Arg is associated with adult atopic dermatitis in Japan.";
RL Br. J. Dermatol. 142:1003-1006(2000).
RN [11]
RP VARIANT PRO-786.
RX MEDLINE-21405389; PubMed-11513543;
RA Andrews R.P., Burrell L., Rosa-Rosa L., Cunningham C.M., Brzezinski J.L., Bernstein J.A., Khurana Hershey G.K.;
RT "Analysis of the Ser786Pro interleukin 4 receptor alpha allelic variant in allergic and nonallergic asthma and its functional consequences.";
RL Clin. Immunol. 100:298-304(2001).
RN [12]
RP VARIANT ILE-579.
RX MEDLINE-21182061; PubMed-11285129;
RA Lozano F., Placés L., Vila J.M., Padilla O., Arman M., Gimferrer I., Sazano B., Lopez de la Igulesia A., Miserachs N., Vives J.;
RT "Identification of a novel single-nucleotide polymorphism (Val554Ile) and definition of eight common alleles for human IL4RA exon 11.";
RL Tissue Antigens 57:216-220(2001).
CC -1- FUNCTION: THIS IS A RECEPTOR FOR INTERLEUKIN-4. A SOLUBLE FORM OF THE IL-4 RECEPTOR MAY REPRESENT A REGULATORY MOLECULE SPECIFIC FOR IL-4-DEPENDENT IMMUNE RESPONSES.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA CHAIN AND A COMMON GAMMA CHAIN.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DISEASE: DEFECTS IN IL4R ARE A CAUSE OF ATOPIC ASTHMA.
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC -1- DATABASE: NAME=PRO; NOTE=CD guide CD124 entry;
CC WWW=<http://www.ncbi.nlm.nih.gov/prov/cd/cd124.htm>
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X52425; CA36672.1; -
CC EMBL: AC004525; AAC23495.1; -
CC PIR: A60386; A60386.
CC PIR: A47603; A47603.
CC MIM: 147781; -
CC InterPro: IPR002996; CRIA.

```

DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003931; Hematopo_receptor_S_F1.
DR Pfam; PF00041; fn3; 1.
DR SMART; SM00060; FN3; 1.
DR PROSITE; PS01355; HEMATOPO_REC_S_F1; 1.
KW Receptor; Transmembrane; Glycoprotein; Signal; Disease mutation;
KW Polymorphism.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 825 INTERLEUKIN-4 RECEPTOR ALPHA CHAIN.
FT DOMAIN 26 825 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 233 256 POTENTIAL.
FT DOMAIN 257 825 CYTOPLASMIC (POTENTIAL).
FT DISULFID 34 44 BY SIMILARITY.
FT DISULFID 74 86 BY SIMILARITY.
FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 75 75 I -> V (IN ATOPIC ASTHMA).
FT VARIANT 400 400 /FTIG=VAR_008034.
FT VARIANT 400 400 E -> A (IN DBSNP:1805011).
FT VARIANT 431 431 /FTIG=VAR_011657.
FT VARIANT 431 431 C -> R (IN DBSNP:1805012).
FT VARIANT 436 436 /FTIG=VAR_011658.
FT VARIANT 436 436 S -> L (IN DBSNP:1805013).
FT VARIANT 503 503 /FTIG=VAR_011659.
FT VARIANT 503 503 S -> P (IN DBSNP:1805015).
FT VARIANT 576 576 /FTIG=VAR_011660.
FT VARIANT 576 576 Q -> R (IN ATOPIC PATIENTS).
FT VARIANT 579 579 /FTIG=VAR_008035.
FT VARIANT 579 579 V -> I.
FT VARIANT 579 579 /FTIG=VAR_011661.
FT VARIANT 752 752 S -> A (IN DBSNP:1805016).
FT VARIANT 752 752 /FTIG=VAR_011662.
FT VARIANT 786 786 S -> P (IN 1.8% OF THE POPULATION; IN DBSNP:1805014).
FT VARIANT 786 786 /FTIG=VAR_011663.
FT SEQUENCE 825 AA; 89658 MW; 9F886DF5612297F8 CRC64;

Query Match 2.9%; Score 107; DB 1; Length 825;
Best Local Similarity 19.0%; Pred. No. 1.3;
Matches 154; Conservative 75; Mismatches 270; Indels 312; Gaps 38;

QY 41 LWDSDILPLGDIIVP-----APGPVLAETHLOTLEVLRCQETDCDCLRLVAVHLAVH 93
DQ 103 LWAGQOLLWKSGFKPSEHYKPRAPGNLTVTHTNVSDTLTWTNSPPYDPDNYLNLHTYAVN 162
QY 94 GHWEPEDEKFGGAADSGVEEPNRLAQAVVLSFOAYPTARCVLLE--VQVPAALVQF 151
DQ 163 -IWEENDPAD-----FRIV---NVTLEPSLRITAASTLKS 193
QY 152 GOSVGSVYDCFEALGSEVRWISYTPRYEKELNHTQQLPALPWLNVSDGDNVHLVLN 211
DQ 194 GTS-----YRARVRAWAQC-----NTTWSEWSPKTKWNSYREPEQHLILG 236
QY 212 VSEEQHFGLSL-----YWNQVQPKPR-----WHNLTGPQI 244
DQ 237 VSVSCIVILAVCLLCYVSTIKKEWMDQIPNPARSLVAIIQDAQSGQWKRSGQEP 296
QY 245 ITLNH-----TDLVPLCLIQWPLEPDSVRTNICPREDPR-----AHQNLWQAA 289
DQ 297 AKCPHWKNCITKLPCF-----LENMKRD-----EDPHKAEMPFQSGSKSAWCPV 344
QY 290 RLRLTLQSW-----LIDAPCSLPAEALCWAPGDPGPOPLVPLSWENVTV 338
DQ 345 EISKIVL--WPESISVVRVCVELFEAPVECEEEV--EEEGSGFC---ASPESRRDQF 397
QY 339 VNSSEKLQIQLWADSLGPKDDVLLLETRG---PDNRSKLCALFESGCTS-----L 388
DQ 398 GREGIVARLTESLFLDLG-----BENGFGCQDDMGESCILLPPSGTSAHMPWDEF 448

```

RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmshurst J., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.,
RT "The DNA sequence and comparative analysis of human chromosome 20.",
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE OF 733-1709 FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Spleen;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
脾 spleen.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: MACROPHAGE-RESTRICTED ADHESION MOLECULE THAT MEDIATES
CC STAILIC-ACID DEPENDENT BINDING TO CELLS.
CC -! SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (ISOFORM 1) AND
CC SOLUBLE (ISOFORM 2).
CC -! ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -! SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -! SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -! SIMILARITY: CONTAINS 16 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -! DATABASE: NAME=PROV; NOTE=C-D guide CD169 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/985165905_g.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; AF230073; AAC00757.1; -
DR EMBL; ALI09804; CAC17543.1; -
DR EMBL; ALI09804; CAC17542.1; -
DR EMBL; AK024462; BAB15752.1; -
DR EMBL; AK024459; BAB15749.1; -
DR EMBL; AK024479; BAB15769.1; -
DR MIM; 600751; -
DR InterPro: IPR003599; IG_MHC.
DR InterPro: IPR003006; IG_IHC.
DR InterPro: IPR003598; IG_C2.
DR InterPro: IPR003600; IG_Like.
DR Pfam; PF00047; Ig_14.
DR SMART; SM00409; IG; 17.
DR SMART; SM00408; IGC2; 14.
DR SMART; SM00410; IGL_Like; 1.
KW Signal; Glycoprotein; Cell adhesion; Immunoglobulin domain; Repeat;
KW Transmembrane; Alternative splicing; Antigen.
FT SIGNAL 1 19
FT CHAIN 20 1709
FT TRANSHEM 1642 1662
FT DOMAIN 1663 1709
FT DOMAIN 20 136
FT DOMAIN 153 224
FT DOMAIN 255 312
FT DOMAIN 339 397
FT DOMAIN 426 498
FT DOMAIN 524 582
FT DOMAIN 617 696
FT DOMAIN 722 781
FT DOMAIN 810 883
FT DOMAIN 909 967
FT DOMAIN 998 1074
FT DOMAIN 1100 1156
FT DOMAIN 1186 1248
FT DOMAIN 1274 1331
FT DOMAIN 1350 1432
FT DOMAIN 1458 1518
FT DOMAIN 1547 1620

FT	DISULFID	41	98	BY SIMILARITY.
FT	DISULFID	160	217	BY SIMILARITY.
FT	DISULFID	262	305	BY SIMILARITY.
FT	DISULFID	346	390	BY SIMILARITY.
FT	DISULFID	433	491	BY SIMILARITY.
FT	DISULFID	531	575	BY SIMILARITY.
FT	DISULFID	624	689	BY SIMILARITY.
FT	DISULFID	729	774	BY SIMILARITY.
FT	DISULFID	817	876	BY SIMILARITY.
FT	DISULFID	916	960	BY SIMILARITY.
FT	DISULFID	1005	1067	BY SIMILARITY.
FT	DISULFID	1193	1241	BY SIMILARITY.
FT	DISULFID	1281	1324	BY SIMILARITY.
FT	DISULFID	1367	1425	BY SIMILARITY.
FT	DISULFID	1465	1511	BY SIMILARITY.
FT	DISULFID	1554	1613	BY SIMILARITY.
FT	CARBOHYD	159	159	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	265	265	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	339	339	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	499	499	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	697	697	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	726	726	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	730	730	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	741	741	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	886	886	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	1104	1104	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	1138	1138	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	1251	1251	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	1462	1462	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	CARBOHYD	1476	1476	N-LINKED (GLCNAC . .) (POTENTIAL).
FT	VANSPIC	1632	1709	ALHRIHORQQLMWTGLVGLLLILLGAGACTWRRRYCK OSMGNSGVMAOKETTDLIPDAATCESTCAPLG -> GSGRDLHPGHSAOKRPSS (IN ISOFORM 2).
FT	CONFLICT	1349	1349	A -> T (IN REF. 1).
FT	CONFLICT	1519	1519	A -> V (IN REF. 3); BAB15749 AND BAB15769).
SQ	SEQUENCE	1709 AA;	182624 MM;	587C/CCA0B789A6D CRC64;
Query Match 2.9%; Score 106.5; DB 1; Length 1709;				
Best Local Similarity 21.1%; Pred. No. 3.6;				
Matches 149; Conservative 68; Mismatches 262; Indels 227; Gaps 33				
OY	62 AFTHLO-TEL-----VLRCORETDCDLCLRAVNHLAVHGHWBPDEKFGGAADS	111	:	:
DB	983 ABRHTLTTLMTGTGFRGLLCRVSDPRADQRL-----LHC-----DELVAStIQ	1029	:	:
OY	112 GVEEPRNASLAQNVLSFOAYPTARCYLEVOVPAAIYO-----FGQSASIV	159	:	:
DB	1030 GVGGEGGSPPRLHVVA-----PNT-----LRLEIHGMLEDEGYICEASNTLQAASASAD	1081	:	:
OY	160 YCCFEALGSEVRIMSYTPRYEKELNHNQOLPALPLWNLVNSADGDVHLVLVNSEQHNG	219	:	:
DB	1082 FD-----ADAIVNQVMFGATVREGQLVN-----LTCILWTTTHPAO-----	1116	:	:
OY	220 ISLYNNQVGPRPKPMWNKNKLTPQITLNHTLDLPCLCIOWPLPEDSVTNIGCFREDP	279	:	:
DB	1117 LLYTYTQ-----DGOQRIDA-HSIRPLPVUTYR-----DATSVRCGCVGRGRAP	1158	:	:
OY	280 RAHQMLKQARLRLLTLTOSWLIDARPSCFLPAEAALCKRARCGDRCPORLVPPLSMENVTVDY	339	:	:
DB	1159 RLS-----RPITLD--VLVARPNRLRLTYLL--ESHGGO-----LALVLCTVDVS	1197	:	:
OY	340 NSSEKLODEC--LMWDSIGPLKDOVYLLETGPQON-----KSICALPESG--CTS----	388	:	:
DB	1198 RPPAOLAISHAGRLLASSTAASVPMTRLRLERLKGPOPRDEGFYSCSARSPLGQAMNSTLEJR	1257	:	:
OY	389 -----PSKASTRAARLEGLEYLLDDLOSGCCLAMDDELALMACPMDKYTIIKKFWALY	439	:	:
DB	1258 LEGVAVILIAPAAYAEGBAPITVTCADPAHANAPPLYTW-----YHNGWM-----	1300	:	:
OY	440 WLACLFFAAISLILLLKRDHAKGWLRLKODVRSAGAARGRAALLLYSADDSGFERLWG	499	:	:


```

Db 1301 ----LOEGPAASLSFLVATRAHAGAYSCQAQDAQGRSSPAALQVLYAPQDAVLSSFRD 1356
QY 500 ALASALCOLPLRVAVDLWSRRSLSAQGPVAVFHAQRROTLEQGGVVLLFSPGAVALCSE 559
Db 1357 SRASWAVTQCTVDSEPPALASHDGKV-----LATSGVHSLASGTHGVQVARN 1407
QY 560 WL-----QDGVSG-----PGAHGPHDAFRASISCV 584
Db 1408 ALRLQVQDVPAGDDTVCTAQNLLSGISITIGRLQVEGARVVAEPGLDVPEGA-ALNLSR 1466
QY 585 LPDFLQGRAP-GSYGACF--DRLLHPDAVPALFRT-----VPVFTLPSQLPDFLQALQ 635
Db 1467 L---LGGPGVGNSTFAWMDNRRLHAEPVPTLAFTHVARAQAGMYHCLAEPL--TGA-- 1519
QY 636 QPRAPRSGRQLQRAEQVSRALOPALDSYFHPGCTPAPGRGVGPCAG 681
Db 1520 -----AASAPVMLRVLYPKTPTMVAEVEPEG 1547

RESULT 9
TLE4_RAT STANDARD; PRT; 741 AA.
AC 007141.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Transducin-like enhancer protein 4 (ESP2 protein).
GN TLE4 OR ESP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Hippocampus;
RX MEDLINE=94064640; PubMed=8245004;
RA Schmidt C.J., Sladek T.E.;
RT "A rat homolog of the Drosophila enhancer of split (groucho) locus
RT lacking WD-40 repeats.";
RL J. Biol. Chem. 268:25681-25686(1993).
CC -!- FUNCTION: NUCLEAR EFFECTOR MOLECULE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT
CC PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L14463; AAC37640.1;
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40. 5.
DR PRINTS; PR00320; GPROTEINBRT.
DR SMART; SM00320; WD40. 6.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS50082; WD_REPEATS_2; 2.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Nuclear protein; Phosphorylation; Repeat; WD repeat.
FT DOMAIN 1 106 GLN-RICH (OR Q DOMAIN).
FT DOMAIN 107 173 GLY/PRO-RICH (OR GP DOMAIN).
FT DOMAIN 175 242 CCN DOMAIN.
FT REPEAT 441 472 WD 1.
FT REPEAT 499 529 WD 2.
FT REPEAT 543 573 WD 3.
FT REPEAT 585 615 WD 4.
FT REPEAT 667 697 WD 5.
FT REPEAT 708 738 WD 6.

```

```

FT MOD_RES 218 218 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD_RES 233 233 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
SQ SEQUENCE 741 AA; 80490 MW; 43F105E568FD87A5 CRC64;

Query Match 2.98; Score 105.5; DB 1; Length 741;
Best Local Similarity 19.3%; Pred. No. 1.5;
Matches 104; Conservative 54; Mismatches 147; Indels 235; Gaps 27;

QY 23 RLVGPDATHSPGLSCLWSDILCLPGDIVPAPG--PVLAP--THLOTELVLRCOKET 78
Db 297 RTDAPPTGSGNSTPGLR-----PVGKPPGVDPLASSURTPMAYPCPPT 340
QY 79 DCDLCLRVAVHLAVHGWEEPEDEKFGGAADSGVEEPRNASLQAQVYVLSFYQAYTARCV 138
Db 341 P-----FGIVPHAGMNGELTSP-----GAAYAGL--HNISPMQMSAAA----- 376
QY 139 LLEVQVPAALVQFGSGVGVYDCFEAALGSEVRISWSTQPRYKELNHTQOLPALP--- 195
Db 377 -----AAAAAAYGRS-----FVVGFDPHHMRVPAIPPNL 407
QY 196 -----WLVNSADGDNVHLVNLVNSEEQHFGLSLYWNOVGPPKPRHKNLTGP-- 242
Db 408 TGIPGKPAYSFHVSADG-----QNPVFPF--PDALLGPGIP 443
QY 243 ----QIITLNHTDLVPCL-----CIQVWPLEPDSVYRTNICPFREDPRAHQ 283
Db 444 RHAQINTLNHGEVVCATISNPTRHVYTGKGVKVDITDPGNKSPVSQL--DCLNRD 501
QY 284 NLWQAARL-----RL-----TLQSWLLDAP-----CSLPAEALCHWRAGGDCQPLV 327
Db 502 NYIRSCRLLPDGRTLVGGGEATLSIWDLAAPTPIKAELTSSAPACY----- 549
QY 328 PPLSWENVTVDVNSSEKILQLEQ-----LWADSLGPLKDDVLLLETRGPQDNRSLSALE 381
Db 550 -----ALAISSPSKVCFCSCSDGNIAW-----DLHQTIVKQFGHTDGCIDIS 596
QY 382 PSQGTSLPSKASTRAARLGEYLLDQSGQCQLQWD--DDLGLAWACPM----- 428
Db 597 NDG-TKLWTGCLDNTVR-----SWDLREGLOLQHDFTSQIFSLGVCYCTGEWLAVGMENS 650
QY 429 -----DKY---IHKRWALVWLACLLFAAALSLILLKKHAKWLRLLKODVRSRG 475
Db 651 NVEVLHVTKPKDKYQLHLHE-----SCVLSLKFAHCGKWF-----VRPG 688

RESULT 10
TLE4_HUMAN STANDARD; PRT; 766 AA.
AC Q04727; Q9ULF9; Q9NSL3;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transducin-like enhancer protein 4.
GN TLE4 OR KIAA1261.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
RN [2]
RP SEQUENCE OF 44-766 FROM N.A.
RC TISSUE-Brain;
RA Bloeker H., Boecher M., Brandt P., Meves H.-W., Weil B., Wiemann S.;

```

Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.

[3]

Sequence of 319-766 FROM N.A.

TISSUE-Fetal brain;

Medlin=93265135; PubMed=1303260;

Stifani S., Blumheller C.M., Redhead N.J., Hill R.E.,

Artavanis-Tsakonas S.;

"Human homologs of a Drosophila Enhancer of split gene product define

a novel family of nuclear proteins.";

Nat. Genet. 2:119-127(1992).

CC -1- FUNCTION: NUCLEAR EFFECTOR MOLECULE.

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -1- TISSUE SPECIFICITY: IN ALL TISSUES EXAMINED, MOSTLY IN BRAIN, AND

MUSCLE.

CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).

CC -1- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT

PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation-

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib-sib.ch).

DR EMBL; AB033087; BAA86575.1; ALT-INIT.

DR EMBL; AL162059; CAB82397.1; .

DR EMBL; M9439; AAA61195.1; .

DR MIM; 605132; .

DR InterPro; IPR001680; WD40.

DR Pfam; PF00400; WD40; 12.

DR PRINTS; PR00320; GPROTEINRPT.

DR SMART; SMO0320; WD40; 7.

DR PROSITE; PS00678; WD_REPEATS_1; 2.

DR PROSITE; PS0082; WD_REPEATS_2; 3.

DR PROSITE; PS0294; WD_REPEATS_REGION; 2.

FW Nuclear protein; Phosphorylation; Repeat; WD repeat.

FT DOMAIN 1 130 GIN-RICH (OR Q DOMAIN).

FT DOMAIN 200 267 GIN/PRO-RICH (OR GP DOMAIN).

FT REPEAT 466 497 WD 1.

FT REPEAT 524 554 WD 2.

FT REPEAT 568 598 WD 3.

FT REPEAT 610 640 WD 4.

FT REPEAT 692 722 WD 5.

FT REPEAT 733 763 WD 6.

FT DOMAIN 398 408 POLY-ALA.

FT CONFLICT 125 125 Q -> QQ (IN REF. 2).

FT CONFLICT 461 461 A -> P (IN REF. 3).

FT CONFLICT 502 502 C -> A (IN REF. 3).

SO SEQUENCE 766 AA; 82911 MW; 7CB251E11918EE CRC64;

Query Match 2.8%; Score 105; DB 1; Length 766;

Best Local Similarity 19.2%; Pred. No. 1.7;

Matches 96; Conservative 58; Mismatches 148; Indels 198; Gaps 24;

QY 23 RLVGQDATHCSPGLSCRLMDSIDILCPDIYPAPG-FYLAP-THLDTELVLRCKEY 78

DB 322 RFDAPTPGNSSTPGLR-----FVPCRPQGVNRLASSLRTPMVVPCEPT 365

QY 79 DCDLRLVAVHLAVGHWEPEDEKFGAGADSGVEEPNNAQLQAVVSFGAYTPARCV 138

DB 366 P---FGIYPAHAGMNGELTSP-----GATYAGL--HNISQMSAAA 401

QY 139 LLEVQPALVQSGSVGVYDCEFAALGSEVRIMSYTPQRYEKLHNTQOLPALP--- 195

DB 402 -----AAAAAAAYGRS-----RVVGFDRNNHMRVALRPNL 432

QY 136 -----WLVNSADGNVHLVNLVNSEQHNGLSLYNQVOGPKPKNNKVLGP--- 242

DB 433 TGIPGGRPAVSPFHVASDG-----QKQPVDFP--PDALIGFGIP 468

QY 243 -----QITLNHTDLVPLC-----CLOWMPLDPSVRINICPFREDPRAHQ 283

DB 469 RHARQINTLNHGEVAVCAVTSNPRHYTGKGCYKWDISHPCNKSVSOL--DCLNRD 526

QY 284 NLMQARL-----RL-----TLQSWLLDAP-----CSLPAEALCMRARGDPCQPLV 327

DB 527 NYTRSCRLPDGRLIYIGGEASTLSINDLAPTRIRAEITSSAPACY----- 574

QY 328 PPLSMENVTVDNSEKRLQEOC-----LWADSLGPLKDDVLLLETRPGDNRLSCALE 381

DB 575 -----ALAIKSPDSKVCFCSCSDGNLAW-----DLHNGTLVROFGHTDGASCIDIS 621

QY 382 PSCGTSLPSKASTRAALGELYLDLQSGCQLMD--DDGALMAGPMOKYIKRNALV 439

DB 622 NDG-TKLMTGGLDNTVR-----SWDLREGROLQOHDFTSQIFSGYCPTEG----- 666

QY 440 WLACILFAALSLILLKRD 459

DB 667 WLAVGMENSVEVLHTKPD 686

RESULT 11

ID RCQ4_HUMAN STANDARD; PRT; 1208 AA.

AC 094761;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE ATP-dependent DNA helicase Q4 (RecQ protein-like 4).

GN RCQ4 OR RECQ4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Testis;

RX MEDLINE=99097344; PubMed=9878247;

RA Kitao S., Ohsu I., Ichikawa K., Goto M., Furuchi Y., Shimamoto A.;

RT "Cloning of two new human helicase genes of the RecQ family:

RT biological significance of multiple species in higher eukaryotes.";

RL Genomics 54:443-452(1998).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=20021764; PubMed=10552928;

RA Kitao S., Lindor N.M., Shiratori M., Furuchi Y., Shimamoto A.;

RT "Rotmund-Thomson syndrome responsible gene, RECQL4: genomic structure

RT and products.";

RL Genomics 61:268-276(1999).

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -1- DISEASE: DEFECTS IN RECQL4 ARE A CAUSE OF ROTMUND-THOMSON

CC SYNDROME (RTS). A DISEASE CHARACTERIZED BY DERMATOLOGICAL FEATURES

CC SUCH AS ATROPHY, PIGMENTATION, AND TELANGIECTASIA AND FREQUENTLY

CC ACCOMPANIED BY JUVENILE CATARACT, SADDLE NOSE, CONGENITAL BONE

CC DEFECTS, DISTURBANCES OF HAIR GROWTH, AND HYPOGONADISM.

CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. RECQ SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation-

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib-sib.ch).

DR EMBL; AB006532; BAA74453.1; .

DR EMBL; AB026546; BAA86899.1; .

DR MIM; 603780; .

DR MIM; 268400; .

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR001878; znf_CCNC.

```
DR Pfam: PF00270; DEAD: 1.
DR Pfam: PF00271; helicase_C: 1.
DR SMART; SM00487; DEXDC: 1.
DR SMART; SM00490; HELIC: 1.
DR SMART; SM00343; Znf_C2HC: 1.
KW Hydrolase; Helicase; ATP-binding; Nuclear protein.
FT NP_BIND 502 509 ATP (POTENTIAL).
FT SITE 605 608 DEAD BOX.
SQ SEQUENCE 1208 AA; 133076 MW; CB809A7765AB48A1 CRC64;

Query Match
Best Local Similarity 2.8%; Score 102; DB 1; Length 1208;
Matches 111; Conservative 72; Mismatches 182; Indels 222; Gaps 31;

QY 16 PVLVSLERLVGPQDATHCSPLGSC---RLWDSIL---CLPGDIVPAPGPVLAPTHLO--- 67
DB 619 PCVLRVCKVLRRMGVHCFGLGTATATRTASDAQHLVAEERPDGLGAPVPVNLHLSV 678
QY 68 -----TELVLRCQKETDCDLCRLVAVHL--AVHGHWEPEDEE 103
DB 679 SMDRDTQALLTLLOGKRFONLDSIIYCNRRDTE---RIAALLRTCLHAAWVPGS--- 732
QY 104 KFGGAADSGVEEPRNASL-----QAQVLSFOAY-----PTARCVLLEV 142
DB 733 --GGRAPKTTAEYHAGMSRRRRVQORAFMQGLRVVVAVATVAFGMGLDRPDVRVAV-LHL 789
QY 143 QVPAALVQSGSYGVVYDFCAALGSEVRISYTPRYE--KELNHTQQLPALPWLNV 200
DB 790 GLPPSFESIVQAVGRAGD-----QPAHCHLFLQPGEDLRLRHVHADSTDFLAVK 843
QY 201 AGDGNVHLVNV-----SEEHFGLSLYNQVQGP-PKPRW-----HKN 238
DB 844 -----RLQVRFPACTCTCTRPPEQE-----GAVGGERPVKYPQEAQLSHQA 889
QY 239 LTGPQIITLNTDLVPC-LCIQWPLEPSVRTNICFPREDPRAHQLWQARLLRLL--T 295
DB 890 APGPRVCMGHERALPIQTQVQALDMPPEAIETLLCYL-----ELHPPHW-----LELLATT 941
QY 296 LQSWLLDAPCSLPAEALCWAP-----GGDPQC-----PLVPLPSWENV 335
DB 942 YTHCLNCPGPGQAQQLAHLRCPPLAVCLAQPLPDPGQSSSVFDMVLVDVSGWELA 1001
QY 336 TV-----LQDLQS---GQC-LQLWDDDLGALWACPMDKYIHR 435
DB 1002 SVRRALCQLQWDHEPRTVGRRGTGLVEFSELAFLHRSFGDLTAEEKDQICDFLY----G 1057
QY 358 PLKDDVLLLETGPPQDNRSISCALE-----PSGCTSLPSKASTRAAR-----LGEY 402
DB 1058 RVQ-----ARERQALRLRRTTQAFHSAFFSPCGPCLEQDDEERSTRKLDLGRY 1107
QY 403 L-----LQDLQS---GQC-LQLWDDDLGALWACPMDKYIHR 435
DB 1108 FEEEGQEPGGMEDAGPEPGARLQWEDQV-----RCDIRQFLSLR 1150

RESULT 12
NER3_RAT
ID NER3_RAT STANDARD; PRT; 418 AA.
AC Q99PW5;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (Ganglioside sialidase)
DE (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
```

```
RX MEDLINE-21092669; PubMed-11162581;
RA Hasegawa T., Feijoo Carnero C., Wada T., Itoyama Y., Miyagi T.;
RT "Differential expression of three sialidase genes in rat
RL Biochem. Biophys. Res. Commun. 280:726-732(2001).
CC -|- FUNCTION: Plays a role in modulating the ganglioside content of
CC the lipid bilayer at the level of membrane-bound sialyl
CC glycoconjugates (By similarity).
CC -|- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates
CC -|- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -|- TISSUE SPECIFICITY: Expressed in brain, cardiac muscle and weakly
CC in liver.
CC -|- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.
CC -|- SIMILARITY: CONTAINS 3 BNR REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB026841; BAB32440.1;
DR InterPro; IPR002860; BNR.
DR Pfam; PF02012; BNR. 3.
KW Hydrolase; Glycosidase; Membrane; Repeat.
FT REPEAT 129 140 BNR 1.
FT REPEAT 201 212 BNR 2.
FT REPEAT 252 263 BNR 3.
FT SITE 24 27 FRIP MOTIF.
FT ACT_SITE 25 25 BY SIMILARITY.
FT ACT_SITE 45 45 POTENTIAL.
FT ACT_SITE 50 50 POTENTIAL.
FT ACT_SITE 87 87 POTENTIAL.
FT ACT_SITE 223 223 POTENTIAL.
FT ACT_SITE 243 243 POTENTIAL.
FT ACT_SITE 339 339 BY SIMILARITY.
FT ACT_SITE 369 369 POTENTIAL.
FT ACT_SITE 386 386 POTENTIAL.
SQ SEQUENCE 418 AA; 46980 MW; 7CC46F2E5952E240 CRC64;

Query Match
Best Local Similarity 2.7%; Score 101.5; DB 1; Length 418;
Matches 106; Conservative 43; Mismatches 135; Indels 175; Gaps 29;

QY 182 EKELNHTQQLPALPWL-----NVSADGDNVHLVNSEQHFGLSLYNQVQ 228
DB 17 EQKNRITVRIPALLYIPPTHFLFAEMRTSRDEDAVILVFRGVMK--GCSVW---- 70
QY 229 GPPKPRWHKNTGPOIITLNTDLVPCLCIQWPLEPSVRTN---ICFPREDPRAHQL 285
DB 71 GPQQLMEATLPG-----HRTMSPC---PW--EKNTGRVYLFFICVQGHVSERWQL 118
QY 286 W--QAARLRLLTLOSLLDAPCS-----DSGCSNGEVDLTEEVVSGEMKHWATFVGFHGIQLQSGRL 174
DB 119 WGRNAARLCFLYSE-----DSGCSNGEVDLTEEVVSGEMKHWATFVGFHGIQLQSGRL 174
QY 307 -LPREAL--CWRAPGDPQPL-VPPLSW-----ENVTVDNSSKQLQLO---ECLWADS 355
DB 175 LIPAYAYLISCWFL-----CFPCSVKPHSLMFYSDDLGVTWCHGKRFKIPQVTGECQAEV 229
QY 356 LGPLKDDVLLLETGPPQDNRSISCALE-----NRSLEALEP--SGCTSLPSKASTR 395
DB 230 PKKAGNVLYCSARTPNKFAEAFSTDSGDCFQKPTLQOQLC--EPRGCGOQ--SVWSTR 285
QY 396 AARLGEYLLQDSQCQLQWLWDDDLGALWACP-MDKYIHRKRWLVWLACLLFAAALSIL 454
DB 286 PLKN-PYTCQD--SSGK-----DVPSTQKCPCLMDRSL----- 314
```

```

OY 455 LKKHAKGWLRLKODVSGAARGRAALLYSADSGFERLVGALASALCOLPLXAV 514
DB 315 -----EVEGAGAPS-GTWLYSHPTNKKRM--NLGIYNNPBLE--V 353
OY 515 DLSRRELSSAGPVAMFAORQTLQEG-GVYVLLFSPG 552
DB 354 NWSRPMIIRKGPSTGSDL-----AVVEGOGCLFACLEPCG 388

RESULT 13
M3K6_HUMAN
ID M3K6_HUMAN STANDARD: PRT; 1011 AA.
AC O95382;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase 6 (EC 2.7.1.-).
GN MAPKK6 OR MAPKK6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99092374; PubMed=9875215;
RA Wang X.S., Diener K., Tan T.-H., Yao Z.;
RT "MAPKK6, a novel mitogen-activated protein kinase kinase, that
RT associates with MAPKK5."
RL Biochem. Biophys. Res. Commun. 253:33-37(1998).
CC -!- FUNCTION: ACTIVATES THE JNK, BUT NOT ERK OR P38 KINASE PATHWAYS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL: AF100318; AAD05304.1; -.
DR MIM: 604468; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00220; S_TKC.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS00108; PROTEIN_KINASE_SF.1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM.1.
DR Transferrase: Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 371 629
FT NP_BIND 377 384 ATP (BY SIMILARITY).
FT BINDING 400 400 ATP (BY SIMILARITY).
FT ACT_SITE 494 494 BY SIMILARITY.
SQ SEQUENCE 1011 AA; 112512 MW; F7294BA0587D2EC9 CMC64;

Query Match 2.7%; Score 100; DB 1; Length 1011;
Best Local Similarity 21.2%; Pred. No. 6;
Matches 149; Conservative 91; Mismatches 268; Indels 194; Gaps 38;

```

```

OY 219 -GLSL-----YANVOGPPKPRMHNKILGPOIILNHTDLVPC-----L 256
DB 141 IGMKGLCLLARKGVCKMYTYWD-----GFYLAQILANDPTQVLAIEQLYKL 190
OY 257 CIOVWPL-----EPDSVRNTICPFREDP-----RAHONLM-----QAARL 291
DB 191 NAPIWLVSVHETFLLYHFRPTPEPPGPPRRRAH--FWLHLLDSCPPFTACAGSQOC 248
OY 292 RLTLQSWLIDAPCSLPAEALCWRAPGDCOP-----LYP-----PLSMENVTVDNS 341
DB 249 LVLVLEMKVLLPAKLEVR-----GTDPSVTYTLISLEPEQDIDISSW---TFPVAS 297
OY 342 -----SEKIQDPCIMADSLGPKADVLLLETRGPQDNLSLALPESGCT-----SLPSKAS 393
DB 298 ICGVASAKRDECCFLVLPADVOLCEFPVSG--HCOMFGCLDAWVNTDSTAPABEA 355
OY 394 TRARILGYLLQDLOGGOCLOLMDDDLALNACPMADYIHKRWALWVL--ACLLFPAAL- 450
DB 356 EGAGEMLEFDYEYETGTGRLVKGKTYGVYA--GDRRTFRVIAIKETIPERDSRFSOPLH 414
OY 451 SLILLKKDHAKGWLRLKODVSGAARGRAALLYSADSGFERLVGALASALCOL- 508
DB 415 EELHLRRLRKKNIVRYL-----GSASQG--GYLKFMEVEVG-----GLSLILRSYWG 462
OY 509 PLR---VAVIDMSRRELSAGCPVAMFAORQTLQEGGVYVLLSP-----G 552
DB 463 PLKNESTISPYTROILOGLYLDHNDHIVHND-IGDGVVLIINTFGLKISDFGSKRIA 521
OY 553 AVALCSEWLQGV-----SGPAHGPHDAFPAASISCVLPDLOGRAP-----GSYVG 599
DB 522 GTPCTETFTFTLOVMAPELIDQGRGYG-KAADIWSLACTYIEMATRPPEHLSQOA 580
OY 600 ACPD---RLHPADVPALEFVPTVPSQLPD-----FLGALQOPRAPRS 642
DB 581 AMFQVMKVKVHPMSSLSAEQAFLRTFEPDPLRASAOQLLGDPLQGGKSRSS 640
OY 643 GRLOER-AEYVSRLQPALDSYF-----HPPGTP 670
DB 641 PRHAPRSDAPSASTPSANSTQSQTPPCQAPSOHPPSP 682

RESULT 14
CXAA_MOUSE
ID CXAA_MOUSE STANDARD: PRT; 505 AA.
AC O9WS4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gap junction alpha-10 protein (Connexin 37) (Cx37).
GN GJA10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99262620; PubMed=10329667;
RA Manthey D., Bukauskas F., Kozak C., Willecke K., Lee C.G.;
RT "Molecular cloning and functional expression of the mouse gap junction
RT gene connexin-37 in human HeLa cells."
RL J. Biol. Chem. 274:14716-14723(1999).
CC -!- FUNCTION: ONE GAP JUNCTION CONSISTS OF A CLUSTER OF CLOSELY PACKED
CC PAIRS OF TRANSMEMBRANE CHANNELS, THE CONNEXONS, THROUGH WHICH
CC MATERIALS OF LOW MW DIFFUSE FROM ONE CELL TO A NEIGHBORING CELL.
CC -!- SUBUNIT: A CONNEXON IS COMPOSED OF A HEXAMER OF CONNEXINS.
CC -!- TISSUE SPECIFICITY: LOW LEVELS WERE DETECTED IN SKIN, HEART,
CC KIDNEY, TESTIS, OVARY, INTESTINE. EXPRESSION NOT DETECTED IN
CC BRAIN, SCIATIC NERVE OR LIVER.
CC -!- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. ALPHA-TYPE (GROUP II).
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AJ010741; CAB40358.1; -
 DR MGD; MG1:1339969; Gjal0.
 DR InterPro: IPR000500; Connexin.
 DR Pfam: PF00029; connexin; 1.
 DR PRINTS: PR00206; CONNEXIN.
 DR SMART: SM00037; CNX; 1.
 DR PROSITE: PS00407; CONNEXINS_1; FALSE_NEG.
 DR PROSITE: PS00408; CONNEXINS_2; 1.
 KW Gap junction; Transmembrane.
 FT DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 20 40 POTENTIAL.
 FT DOMAIN 41 76 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 77 97 POTENTIAL.
 FT DOMAIN 98 165 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 166 186 POTENTIAL.
 FT DOMAIN 187 209 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 210 230 POTENTIAL.
 FT DOMAIN 231 505 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 505 AA; 57115 MW; 77FB9575FD0A274A CRC64;

Query Match 2.7%; Score 99; DB 1; Length 505;
 Best Local Similarity 19.6%; Pred. No. 2.9;
 Matches 70; Conservative 50; Mismatches 107; Indels 130; Gaps 20;

QY 125 VVLSFOAYPTARCVLLEVOVPAALVQSGVGVYDCFEALGSEVRWISYTPRYEKE 184
 DB 177 ILYGFQHPYIKC---TQACP-----NSV---DCF-----VSRP----- 205
 QY 185 LNHTQOLPALPWLNVSGADGNVHLVNSEQHFGLSLYWNQVQPPKPRWHKLTGPQI 244
 DB 206 ---TEKTFIMLFHSA---AISLLNILEIFHGIKIRMLADGKSSGNTENETGPPF 259
 QY 245 IYLNHTDVLVCLCIQWVP-----LEPDS-----VRTNCPREDPRAHONLWQAARLLT 295
 DB 260 HSTNYSGTQCMVCSLPERISLQANNKQOIVRN1-----PRS-KSMWQIPHPRQLE 312
 QY 296 L-----QSWL-----LDAPCSLPAEALCNRAPGDCPCQLVPPL----- 330
 DB 313 VDVSCKRQWAKEIESCTQLHVSPC---PHDRSARIHQPCQCHSVFGFANMSQSWFG 370
 QY 331 -----SWENVTVVNSSEKQLQEC-----LWADSLGP--LKD 361
 DB 371 TMTASQHRPSSALETWERSQGPPEASGRSLTDQSHFGSGSARESGVWTDRLGPGSRKA 430
 QY 362 DVL--LLETRGPQ-----DNRSICALPEPSC-----TSLPSKASTRAARLGE 401
 DB 431 SFLSRLMSEKQRHSDSGSSRL-----NSSCLDFSHGENSPSLPSATGHRASVMVK 483

RESULT 15
 W70T_MOUSE STANDARD; PRT; 922 AA.
 AC Q9D2V7;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE 70 kDa WD-repeat tumor rejection antigen homolog.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Kidney;
 RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberis P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -I- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
 CC -I- SIMILARITY: BELONGS TO THE CORONIN FAMILY OF WD-REPEAT PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AK018739; BAB31380.1; -
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 6.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PRODOM: PD000018; WD40; 2.
 DR SMART: SM00320; WD40; 6.
 DR PROSITE: PS00678; WD_REPEATS_1; 1.
 DR PROSITE: PS50082; WD_REPEATS_2; 4.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 2.
 KW Repeat; WD repeat.
 FT REPEAT 75 115 WD 1.
 FT REPEAT 124 163 WD 2.
 FT REPEAT 166 205 WD 3.
 FT REPEAT 209 253 WD 4.
 FT REPEAT 253 295 WD 5.
 FT REPEAT 295 337 WD 6.
 FT REPEAT 337 379 WD 7.
 FT REPEAT 379 421 WD 8.
 SQ SEQUENCE 922 AA; 100860 MW; C8E4BC1EE9CFE242 CRC64;
 Query Match 2.7%; Score 99; DB 1; Length 922;
 Best Local Similarity 20.3%; Pred. No. 6.4;
 Matches 99; Conservative 48; Mismatches 120; Indels 220; Gaps 23;
 QY 12 LGRSPVLSRLVCPQ-DATHCSPGLSCLRLWSDILCLPGDIVPAPGPVLAPTHQLTEL 70
 DB 451 IGTSPSRSLQSLGLPGSCFRHTQGSLLHR---DSHTNLKLNLTTPG----- 496
 QY 71 VLRCOKETD--CDLCLRVAVHLAVHGHWEPEDEEKFAGGADSGVEEPRNASLQAQVVL 128
 DB 497 -----ESDGFANRLRVAVPLSSG----- 516
 QY 129 FOAYPTARCVLLEVOVPAAL-----VOFGSGSVGVYDCFE-----AALGSEVRWIS 175
 DB 517 -----GEVAVLELQKPGRLPDTALTPLQNGTAVMDLVMDPDHRLAVAGEDARILWR 570
 QY 176 YTPRYEKE-----NHTQOLPALPWLNVSD--GDNVHLVNV----- 212
 DB 571 VPPGLENVLTPTPETVLTGHTKEIYSLRPHPLAADVLAASSYDYLTVRWDLTGAEUKL 630
 QY 213 --SEQHFGLSLYWNQVQPPKPRWHKLTGPQIITLNLHTDLVPLCLCQVWPLEPDSVRT 270

```
Db 631 QHQDQIF--SLAWSP-----DGKQLATV-----CKDGHVRVYEP---RS 665
Qy 271 NICPFREDPRAHQNLWQAARLLTLQSWLLDAPCSLPAEAALCWAPGGDPCQPLVPPL 330
Db 666 SPLPLQEGGPEGG--RGARI-----VWVCDGGCLLV-----696
Qy 331 SWENVTVDNSEKIQEQECLWADSLGPKDDVLLLETRGPQDNRSICALEPSGCTSLPS 390
Db 697 -----GFDSEKQLOL-----YIADAL-----AOGPS---ALLGLDVAPOSTLLPS 734
Qy 391 -----KASTRAARLGEYL-----LQDLQ 408
Db 735 YDPDTGLVLLTKGDTVPF-LYEVLPAPFFLECNSTSPDPHKGFVLLPKTECDIQDVE 793
Qy 409 SQCQLQ 415
Db 794 FARCLRL 800
```

Search completed: May 19, 2002, 17:01:06
Job time: 10247 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 14:11:33 ; Search time 63.07 Seconds
(without alignments)
1048.192 Million cell updates/sec

Title: US-09-608-918-10
Perfect score: 3692
Sequence: 1 MPVFWFLSLALGRSPVLS.....TPAPRGVGPAGPGADGT 688
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	126	3.4	3707	2 S18252	heparan sulfate pr
2	125.5	3.4	1366	2 T35985	probable large pr
3	124.5	3.4	761	2 S20458	pqgf protein - Kle
4	120.5	3.3	478	2 D75564	hypothetical prote
5	115.5	3.1	698	2 A82593	hypothetical prote
6	114.5	3.1	1711	1 A55148	protein-tyrosine-p
7	112.5	3.0	1193	2 F83264	hypothetical prote
8	111	3.0	1118	1 A49724	protein-tyrosine-p
9	109.5	3.0	2055	2 T00093	hypothetical prote
10	108.5	2.9	753	1 WZBE8	gene 56 protein -
11	107.5	2.9	433	2 B75376	sensor histidine k
12	107	2.9	825	1 A60386	interleukin-4 rece
13	106.5	2.9	274	2 A83583	probable biotin sy
14	105.5	2.9	679	2 B75262	conserved hypothet
15	105.5	2.9	741	2 B49555	enhancer of split
16	105	2.8	724	2 T47149	hypothetical prote
17	105	2.8	3670	2 T36249	CDA peptide synthe
18	104.5	2.8	238	2 T07041	hypothetical prote
19	103.5	2.8	1085	2 T03531	cobN protein homol
20	102	2.8	2403	2 A59386	sanko - human
21	102	2.8	2629	2 T30987	telomerase-associ
22	101.5	2.7	418	2 JC7588	exo-alpha-stalidas
23	101.5	2.7	859	2 T35785	probable beta-gluc
24	101.5	2.7	1737	2 T00209	MEGF8 protein - hu
25	101	2.7	427	2 E87669	conserved hypothet
26	100.5	2.7	745	2 H85840	hypothetical prote
27	100	2.7	719	2 T35189	probable ATP-depen
28	100	2.7	846	2 T27282	hypothetical prote
29	99.5	2.7	745	2 C99955	hypothetical prote

30	99.5	2.7	1203	2 I55466	N-methyl-D-asparta
31	98.5	2.7	882	2 I38912	receptor tyrosine k
32	98.5	2.7	890	1 A53743	protein-tyrosine k
33	98.5	2.7	1240	2 T30834	nuclear protein SA
34	98.5	2.7	2124	2 T28658	polyketide synthas
35	97.5	2.6	676	1 WMBEX6	uL6 protein - huma
36	97.5	2.6	709	2 F75584	hypothetical prote
37	97.5	2.6	772	2 T30845	hypothetical prote
38	97.5	2.6	850	2 JC5700	Erbb kinase activat
39	97.5	2.6	897	1 A39255	cytokine receptor
40	97	2.6	1288	2 JE0363	mitogen-activated
41	97	2.6	3573	2 S23070	erythronolide synt
42	97	2.6	3739	2 T17410	polyketide synthas
43	96.5	2.6	1309	2 F82207	ATP-dependent helli
44	96.5	2.6	4391	2 A38096	perlecan precursor
45	96	2.6	648	2 T34999	probable neuramida

ALIGNMENTS

RESULT 1
S18252
heparan sulfate proteoglycan - mouse
N:Alternate names: perlecan
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C:Accession: S18252; A31917; B31917; S66460
R:Noonan, D.M.; Fulle, A.; Valente, P.; Cai, S.; Horigan, E.; Sasaki, M.; Yamada, Y.;
J. Biol. Chem. 266, 22939-22947, 1991
A:Title: The complete sequence of perlecan, a basement membrane heparan sulfate prote
adhesion molecule.
A:Reference number: S18252; MUID:92078153
A:Accession: S18252
A:Molecule type: mRNA
A:Residues: 1-3707 <NO>
A:Cross-references: EMBL:W77174; NID:g200295; PIDN:AAA39911.1; PID:g200296
R:Noonan, D.M.; Horigan, E.A.; Ledbetter, S.R.; Vogeli, G.; Sasaki, M.; Yamada, Y.; H;
J. Biol. Chem. 263, 16379-16387, 1988
A:Title: Identification of cDNA clones encoding different domains of the basement mem
A:Reference number: A92680; MUID:89034110
A:Accession: A31917
A:Molecule type: mRNA
A:Residues: 940-1601 <NO2>
A:Cross-references: GB:J04054; NID:g200252; PIDN:AAA39899.1; PID:g200253
A:Accession: B31917
A:Molecule type: mRNA
A:Residues: 1870-2600 <NO3>
A:Cross-references: GB:J04055; NID:g200300; PIDN:AAA39912.1; PID:g200301
R:Schulze, B.; Mann, K.; Battistutta, R.; Wiedemann, H.; Timpi, R.
Eur. J. Biochem. 231, 551-556, 1995
A:Title: Structural properties of recombinant domain III-3 of perlecan containing a g
A:Reference number: S66460; MUID:95377282
A:Accession: S66460
A:Molecule type: protein
A:Residues: 1272-1274, 'X', 1276, 'X', 1278-1279 <SCH>
C:Superfamily: LDL receptor ligand-binding repeat homology; EGF homology; laminin G r
C:Keywords: glycoprotein
F:199-234/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:285-319/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:325-359/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:368-403/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:764-811/Domain: laminin-type EGF-like homology <LEG>
F:1159-1206/Domain: laminin-type EGF-like homology <LEG7>
F:1563-1610/Domain: laminin-type EGF-like homology <EG7>
F:1613-1668/Domain: laminin-type EGF-like homology <LEG8>
F:3163-3198/Domain: EGF homology <EGF>
F:3270-3423/Domain: laminin G repeat homology <LG2>
F:3464-3492/Domain: EGF homology <EGF7>
F:1256,1891,2336,2394,2427/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 3.4%; Score 126; DB 2; Length 3707;


```

0Y 395 RAARLGEVLLOD-----LQSGGCIQIMDDDLGALIMACMGDKYIHKRRNALVWLAJCLLEAAAL 450
Db 262 -----EFLLEADPGSLMAGLRORRLAGDVALNM-----LYDRRLGLGA-LVBSADR 307
0Y 451 SLILLKDKDAKGMRLRLKOV-----RSGAA-----RGRALLLYSADS 492
Db 308 PEEDVROQITTH--WLOALQOQTPPEQOQHYYQLSRRRQALSPLDQLRQRAFGAPCAPPA 364
0Y 493 GFERLVGAL-----ASALCQ-----LPLRVAVDLSRRRELSAOGPVAMFHA 533
Db 365 GFADFCAALQVAPSVSLACQVTPSGEPVATOGFSLPL-----SMWRRRPEEDPALAPAFVP 420
0Y 534 QRRQTL-----QEGGVVYLLFSPG-----AVALCSEWLQDQGV-----G 567
Db 421 QAAAGDLVAKCEKAPLHLHPSPGDPRLRLLRPPFYSPDOAAGLRLGEOLRLFLALRH 480
0Y 568 PGAHGPHDAFAPAS--LSCVLPEDFLOGRAPGSYGACEDRLHLHDA--VPALPFTVAVFTL 623
Db 481 AGGHEHMLFLPGSMQTLQLOPE--PGRPRALLQALIRQLALEVVASLITSP--ESIAIRHL 537
0Y 624 PSQLEDFLQALOQR--APRSGLQERABOVSRAIOLAPDLSYFHPGCPAPG--RGVG 677
Db 538 MAQLEPERLGTSGHQKGMWLAALAGSABDAQMVAROLLS-LITAVVNP-MPAPAPCRGVE 595
0Y 678 PGAGPGAGD 686
Db 596 RLVPYG--GD 603

```

RESULT 4
D75564
hypothetical protein - Deinococcus radiodurans (strain R1)
C.Species: Deinococcus radiodurans
C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C.Accession: D75564
R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shun, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zaleski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A.Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A.Reference number: A75250; MUID:20036896
A.Accession: D75564
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-478 <WHI>
A.Cross-references: GB:AE001869; GB:AE000513; NID:g6457721; PIDN:AAF09658.1; PID:g6457721
A.Experimental source: strain R1
C.Genetics:
A.Gene: DR0061
A.Map position: 1

Query Match	3.3%	Score 120.5;	DB 2;	Length 478;
Best Local Similarity	24.3%;	Pred. No. 0.21;		
Matches 99;	Conservative 39;	Mismatches 148;	Indels 121;	Gaps 17;

```

OY 276 REPRHQNLWQARLRLLTLOSLL-----DAPCSJAPAAALCWAPGDPQCPQLV 327
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 RADPPEPLREAAO----RIRRLARRLRAPQCPGVPAHSVLY---AOEASALS 161

OY 328 PLSMENTVYVNVNSSEKIQLOEC---WADSLGPKDVLLETFRGPQDNRSICALPSPG 384
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 APAAGDGLTLHLGGPDRRLAHHQGLRLSMR-SLGPNNQLVQDETSGPYSGQLALLRP--- 218

OY 385 CTSLSPKASTRAARLGEYLRLDQDSGQCLJLMDDIGALWACPMXYIHKRMALVYLACL 444
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 -DLPASERQLFLVVGGOQLVLFSG----- 242

OY 445 LFAAALSLILLKDKHAKGWLRLKQVRYSGAARGRA-ALLIYSADDSGFERYGALS 503
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 ----DVLRLRRADAQGTHLAR-----LALGAGCALLLPLPEONGRLRLATLAR 290

```

```

QY 504 ALCCQLRLVAVDYLMSRRRLSSAGPRAMVHAQRQTLDGGVVLVLFSPGAV-----554
Db 291 RLKGDPPR-ADDPASGOVAFAQTPTCSLAARR-SLQLEPLRLGRFPAQVHADEAOYAQA 348
QY 555 -----ALCSEMLDQVSGPR-GAHGPHDARASLSCL-----PDFL 589
Db 349 LILHLRPHYAQOYQAEEHAASAELRLREAADVPLRGREAGRPVYLPAIGHPLFVPRLGDPDL 408
QY 590 QGRAGSGYVACFDRLL-LHPDVAFLRFTVPVFLTPSQRLPDFIGAL 634
Db 409 EYRLRG-----DRLTLTRPD-----YRAELVAVLPCQAQAVTQDGL 443

```

```

RESULT      5      hypothetical protein xp2169 [imported] - Xylella fastidiosa (strain 9a5c)
8a2593
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: A82593
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

```

A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82593
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-698 <SIM>
A:Cross-references: GB:AE004030; GB:AE003849; NID:g9107292; PIDN:AAF84968.1; GSPDB:GN
A:Experimental source: strain 95c
R:Simpton, A.J.G.; Relnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
B:lonnes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincanl, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C
., F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2169

Query Match	3.1%;	Score 115.5;	DB 2;	Length 698;
Best Local Similarity	21.8%;	Pred. No. 0.84;		
Matches 160;	Conservative 90;	Mismatches 234;	Indels 251;	Gaps 43;

```

QY 61 LA-----THLOTELYVLRKKE--TDDCLLRVAHVLAHGHHEHDEDEKGG-----G 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 LAPEPELHLOGLLLOHHDHSDAALSRSTR|-----DROGNTTYVMOAHHA 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 108 AADSGVEEPNRNASLOAOVVLVSFOAYP-----TARCVLLEQVVPALVQFGOSV-----G 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 IARGALDEAEQSLSRATRTIIS--DHPLQLAVNATIALRGVDQALTMLSQAVELLPEEP 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 157 SVVYVCCFEAALGSEVRIMSYOPKR-----EK-----ELN--HTQOQLPALMVLNVSADGDN 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 AVLK-----SLG-----FAYLQKGIARAEARAFVETLNHYIVPRLHIAQLAOROG--R 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 206 VHLVNVSEEOHFGSLVWNOVOGPPK---PRMRKNTLGPDIITLNTHTDLPVPCIAIOV-- 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 LDDALRIIEE---ALSL-----PEEDPTALHR--LTGEFELIAGHPDRLHTHLRQVLA 273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 261 -WPLEPDSVRNICEFE-----DRAHONLMQARLRLL----- 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 TWPEDEPRTLEALLTRMKQLOLMDDDARITLDAALDIKPPNH--DLW--IARLAVAPVSGDEAR 321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 294 LTLDOSMLDAPCSLPAAEALCMRAPGDPCCPLPVPPLSMENWVATDVNSSEKLOJQEBCLA 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
Db 332 IVERKLSAMPBHLPALETLM----- 352
Qy 354 DSLGLPKDDVLLLETRGPQDNRSLSALEPSGCTSLPSKASTRAARLGEVLLQ--DLQSGQ 411
Db 353 -SLHDIQNPEAAETVA---RQIVALEPRLSG-----EQRIVEALLQRPPTAVA 399
Qy 412 CLQWDDDLGALWACPMCKYIHKRWAL-VWLACL-----LFAAAALSLILLKKDHAKGWL 465
Db 400 CLQOLIESVPE-----HERTILRWLGLVQDRAGQFEAALATWLQFOAEQAKYRL 449
Qy 466 RLLKQ-----DVRSGAAARGRAALLYSADDSGFRLLYGALA--SALC----- 506
Db 450 PLPPQWTNSKOMPDLAIPAELARPLLINGPPGSHVETVMMMAWASMLCADRYSET 509
Qy 507 -----OLPLURVAVDLWSRRELSAOGP-----VAWFH-----AQR 536
Db 510 PPADPLORYEIVSELTSGTLTPQALIDAW-RQOLFARGIEDGNVIDLWLNDSLLTALR 568
Qy 537 QTLQEGGVVLLFSPGAVLCEWLDQGVSGAGHPHDAFRASLSCLVDFLQGRAPGS 596
Db 569 PHLPEGRLLIIVLRDPRMLL--DWIAYGSPILPLA---LDSLQQAANW-LGDIL-----N 616
Qy 597 YVGACFDRLHP-----DAVPALFRTV-PVFTLPSQLPDLFGALQOQPRAPSGRL 645
Db 617 QIAALHELDLYPHHLIRLDGIEDNPQALATTLEDIFGSPFIPP---SLEAPRLP-AGRW 672
Qy 646 QERAEQSVRALQPAL 660
Db 673 RQYREVLSAFDAVL 687

RESULT 6
A55148
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type OST precursor - rat
N:Alternate names: OST-PRP; osteostesticular protein-tyrosine-phosphatase
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A55148
J:Mauro, L.J.; Olmsted, E.A.; Skrobacz, B.M.; Mourey, R.J.; Davis, A.R.; Dixon, J.E.
R: Biol. Chem. 269, 30659-30667, 1994
A:Title: Identification of a hormonally regulated protein tyrosine phosphatase associated with the rat osteoblast
A:Reference number: A55148; MUID:95074080
A:Accession: A55148
A:Molecule type: mRNA
A:Status: not compared with conceptual translation
A:Residues: 1-1711 <MAU>
A:Cross-references: GB:L36884
C:Comment: The sequence contains ten fibronectin type III repeats and two protein-tyrosine-phosphatase family protein-tyrosine-phosphatase, receptor type OST; fibronectin type III repeat 10; signal sequence; phosphatase; phosphoric monoester hydrolase; receptor; transmembrane protein; 1-18/Domain: signal sequence #status predicted <SIG>
F:19-1711/Product: protein-tyrosine-phosphatase, receptor type OST #status predicted <MAU>
F:1174-1398/Domain: protein-tyrosine-phosphatase, receptor type OST #status predicted <MAU>
F:1350/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1356/Binding site: substrate phosphate (Arg) #status predicted

Query Match 3.1%; Score 114.5; DB 1; Length 1711;
Best Local Similarity 19.8%; Pred. No. 3.1;
Matches 158; Conservative 73; Mismatches 270; Indels 295; Gaps 39;

Qy 5 WFLLSLALG-RSPVLSLERLVGPQDATHCSPLGSLRLWSDILCLPGDIVAPGVPVAP 63
Db 322 WASHKAGLGARDGVVL---KLSGPMESTSLGPEN-----AVFPGP-LPP 364
Qy 64 THLQTELRLCQKCTDCDLCLRLVAVHLAVHGWEEPEDEKFGGAADSGVEE----PRNA 119
Db 365 GHYTLQL-----KVLAGPYDAWVEGSTWLAESA 392
Qy 120 SLQAOVV-----LSQAYPTARCVLLEVOVPAALVOFGSGVS--VYVDFCEAALGS 169
Db 393 ALPREVPGARLWLDGLEASKQPGRRALLYSDAPGSLGNISVPNGATHVIFGCLVPGAHY 452
```

```
Qy 170 EVRIWSTQPRYKELNHTQOLP--ALPWLNVNADGNVHLVNLNVSEQHFGLSLYWNQV 227
Db 453 RVDIASSTGDIDISQISGYTSPLPQSLSEVSRSSPSD-----LTIWGA 497
Qy 228 QGP---PKPRKHKMLT-----GPOIITLNHTDLP--CLCIQWVP---LEPDSV 268
Db 498 PGQLEGYKVTWHQDGSQSPGDLVDLGDPTLSLTLSLVPGSSYTSAWAGNLGSDSQ 557
Qy 269 RTN1CP-----FREDPRAHONLW-----CAARLRLTLQ----- 297
Db 558 KIHSCTRPAPPTNLSLGFHQPAALKASWYHPPGGRDAFHRLRLRLTLTLESEKVLPRE 617
Qy 298 ----SWL-LDAPCSLPAEALCW-----RAPGGDPCQPLVPPLSWENTVDVNSSEK 344
Db 618 AQNFSWAQLTAGCEQVQLSTLWGSERSSSANATGWP--PSAPL-----VNVTSAP 669
Qy 345 LQLOECUWADSLGPKDDVLL-----ETR-----GP-ODNRSLSCALEPS-----GCTSLP 389
Db 670 TOLQ-VSWAHVPGGRSRYQVTLYQESTRTATSIMGPKEDGTSLGLTGTGYKKEVISWA 728
Qy 390 SKASTRAARLGEY-----LLQDLQSGCQLQWDDDLGALWACPMCKYIHKRWALVWL 441
Db 729 GPLYTAANVSAMTYPIPLINELIVSMOAGS-----AVVNL 763
Qy 442 ACILLFAAALSLLILLKKDHAKGWLRLKQDVRSGAAARGRAALLLYSADDSGFRLLVGAL 501
Db 764 A-----W-----PSGLPGGACHAQDSAGHLSWEQ----- 789
Qy 502 ASALCOLPLRVAVDLWSRRELSAOGPVAMFHAQRQRTLOEG-GVVVLLFSPGAVLCEW 560
Db 790 -----PLKLGOELFMLRLDLPCHTISMSVRCRAGPLQASTHLVWLSVEPGPV----- 836
Qy 561 LDQGVSPGAGHPHDAFRASLSCLVLPD-----FLOGRAPGS-----YVGACFDRL 606
Db 837 -EDVLCHP-----EATYALNMTMPAGDNDVCLVVERLVPPGGGTHFVFOVNTSGDALL 889
Qy 607 HPDAVPALEFVTPVFTLPSQLPDLFGALQOQPRAPSGRLQERAFQVSRALQPALDSYFHP 666
Db 890 LPNLMPTTSVRLSLTVL-----GNSRWSRAVSLVCSLSAEAW-HP 929
Qy 667 PGTPAP-----GRGVG 677
Db 930 PELAEPPQVELGTGMG 945

RESULT 7
F83264
hypothetical protein PA3063 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83264
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;
L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337
A:Accession: F83264
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1193 <STO>
A:Cross-references: GB:AE004730; GB:AE004091; NID:g9949154; PIDN:AG06451.1; GSPDB:G1;
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3063

Query Match 3.0%; Score 112.5; DB 2; Length 1193;
Best Local Similarity 23.0%; Pred. No. 2.8;
Matches 155; Conservative 60; Mismatches 276; Indels 183; Gaps 38;

Qy 81 DCLRLVAVHLAVHGWEE-----EPEDEKFGGAADSGV---EEPNRSIQAOVV--- 126
```

Db 293 DLRLNMAAG-RVERAWEVGOELLALRPEDRTLLADLARLGWTCNGPRALGEWQOLLGA 351
Qy 127 -----LSFOAYPTARCVLLEVVQVPAALVQFGSGVSVVYDCFEALGSEVRIW- 174
Db 352 DDPALREHAWRLSLQMFDFSAIELLAPIGAQRQMTDEELDALVYSHETRTGTPGEAWL 411
Qy 175 -SYTOPRYEKE-----LNHTOOLPALP--WLNVSADGNVHLVNVSEHQEGLS 221
Db 412 RGYVO-RYPQRLAWORLOQIILEHTOOLQOEBTGVMAR-----HPLSVKERMQA-E 464
Qy 222 LYWNQVQGPCKPRWKNLGTGQIITLHNTDLVPCLCIQVWPLEPDSVVRTNICPPREDPRA 281
Db 465 THWNLFDP--PROAW-KVLGVDTTRAIREPEFWRLRAALAWALEQD-----DAR- 510
Qy 282 HONLWQAARLLTLTQSWLLDAPCSLPAAALCWAPGDCQPLVPL-SWENVTVDVN 340
Db 511 -----AAYERMLA-----LDRLNSRDEQLTALYRDSNPKQALQVLIGSWORSRDRPR 559
Qy 341 SSEKLQLOECL--WADSLGLKDDVLLLETRGPQDNRSICALPSGCTSLPSKAST----- 394
Db 560 LASALQLAENLHWP-----ALKS--LLAEABG-----LPEAGSPYYW 596
Qy 395 -RAARLGEVLLDQSGQCLOLWDDDLGALWACPMCKY IHKRWALVW-----LAC 443
Db 597 VARARLAE--QEGHGDVAERLYRE--ALVRFPGENLVRE--LLWFYIDRGRDRSLAP 648
Qy 444 LL-----FAAALLILLKKHAKGWRL-LKQDVRSGAAAGRAALL 486
Db 649 LLAQHWGLALRDLTLWLPASA--SELLERNDQALWAFRLYLKSNPNDLVQAYADAL- 705
Qy 487 YSADSGFE-----RLVGALASALCQLPLRVAVDLWS-----RRELSAQGPVAFHQRQTL 539
Db 706 ---DASGYQDKALRLRLRLRLRDREAVRATPDSFATYLRLLAVAOQPLLAQGEARRAWN 762
Qy 540 QEGGVVLLF-----SPGVALCSEWLDQGVSGPAH-GPHDAFRASLSCVLPDFLQ 590
Db 763 GPAMLLQWFEFLDQLAATNQEPLKDNWLA-WARGRLKIGRNEETQAAALS-----Q 815
Qy 591 GRAPGSYVGCDFRLHPDAPALFTVPVFTL---PSQLPDFLQALQPPRPSGRLOE 647
Db 816 NR-----AALQRLERELDELPQAQRVEALVRLGHGGEALGAL-----GDGHSRD 862
Qy 648 RAEQVSRALOPALD 661
Db 863 NREQLRRQAAILLE 876

RESULT 8
A49724
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type H precursor - human
N:Alternate names: protein-tyrosine-phosphatase, stomach cancer-associated type 1; SAP-1
C:Species: Homo sapiens (man)
C>Date: 03-May-1994 #sequence_revision 23-Feb-1996 #text_change 16-Jun-2000
C:Accession: A49724
R:Matcozaki, T.; Suzuki, T.; Uchida, T.; Inazawa, J.; Ariyama, T.; Matsuda, K.; Horita, K.
J. Biol. Chem. 269, 2075-2081, 1994
A:Title: Molecular cloning of a human transmembrane-type protein tyrosine phosphatase an
A:Reference number: A49724; MUID: 94124561
A:Accession: A49724
A:Molecule type: mRNA
A:Residues: 1-1118 <MATO>
A:Cross-references: GB:D15049; NID:q475003; PIDN:BAA03645.1; PID:q475004
C:Genetics:
A:Gene: GDB:PTPRH; SAP-1
A:Cross-references: GDB:305504
A:Map position: 19q13.4-19q13.4
A:Note: highly expressed in colon and pancreatic cancer cells but not in the normal cell
C:Superfamily: protein-tyrosine-phosphatase, receptor type H; fibronectin type III repea
C:Keywords: carcinogenesis; duplication; glycoprotein; phosphoprotein; phosphoric monoes
F:1-27/Domain: signal sequence status predicted <SIG>
F:27-110/Domain: fibronectin type III repeat homology <3FNA>
F:28-1118/Product: protein-tyrosine-phosphatase, receptor type H #status predicted <MAT>

F:28-761/Domain: extracellular #status predicted <EXT>
F:116-199/Domain: fibronectin type III repeat homology <3FNB>
F:203-289/Domain: fibronectin type III repeat homology <3FNC>
F:296-379/Domain: fibronectin type III repeat homology <3FND>
F:385-468/Domain: fibronectin type III repeat homology <3FNE>
F:474-558/Domain: fibronectin type III repeat homology <3FNF>
F:564-658/Domain: fibronectin type III repeat homology <3FNG>
F:667-737/Domain: fibronectin type III repeat homology <3FNH>
F:762-778/Domain: transmembrane #status predicted <TMN>
F:779-1118/Domain: intracellular #status predicted <INT>
F:846-1070/Domain: protein-tyrosine-phosphatase homology <PTPI>
F:35,78,83,107,132,149,172,196,203,286,304,312,329,352,376,383,401,436,439,470,490,55
F:1022/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1028/Binding site: substrate phosphate (Arg) #status predicted

Query Match 3.0%; Score 111; DB 1; Length 1118;
Best Local Similarity 18.4%; Pred. No. 3.4;
Matches 126; Conservative 65; Mismatches 200; Indels 294; Gaps 31;

Qy 56 APGPVLAPTHLOTELVLRCQKETDCDLCLRVAVHLAVHGHWEPEDEEK-----FGGAA 109
Db 387 APNPV-RNLHMET-----QTNSSIALC-----WEVPDGPYQDYTYWVGTYG 427
Qy 110 DSGVEEPRNASLOAQVLSFOAYPTARCVLLEVVQVPAALVQFGSGVSVVYDCFEALG 169
Db 428 DGGGTETRT-----TNTSVTAERLEPGTLYTFS-----VMAEKNRGARS 467
Qy 170 EVRIWYSYTOPRYEKELNHTQOLPALPWLNVNSADGNVHLVNVSEHQEGLSLYWNVOG 229
Db 468 QNVSISIVP-----NAVTSLSKODWNTST-----IALRWTAPOG 502
Qy 230 PKPRW-----HKNLTGPQIITLHNTDLVPCLCIQVWPLEPDSVVRTNICPFREDPRAH 282
Db 503 PQGSSYSVSWVNRGMTDPRQSTSGTD-----ITLKELEAGSL-----YH 544
Qy 283 QNLWQAARLLTLTQSWLLDAPCSLPAAEA-----LCWRAPG----- 319
Db 545 LTVW-AERNEVRGYNSTLTAA--TAPNEVTDLQNETQTKNSVNLWMLWAKPGDPSQLYVW 601
Qy 320 -----GDCQPLVPLSWENVTVDVNS--EKLOQE-----CLWADSLGLKDDVLL 365
Db 602 VOMASKGHPRRGQDPQANWVNTSRTNETWYKVEALEPGTLYNFTWAE-----RNDV-- 654
Qy 366 LETRGPQDNRSICALPSGCTSLPSKASTRAARLGEYLLDQSGQCLOLWDDDLGALWA 425
Db 655 -----ASSTQSLCASTYPTVTITSCVSTAGY-----GVNLIWS 689
Qy 426 CPMCKY-----IHKRW----- 436
Db 690 CPQGGYEAFELEVGQSGQDRSCGEAVSLGLGPARSYPATITIMDKMWSHVVC 749
Qy 437 -----ALVWLACLLFAAALLSLLL-----KKDHAKGWLRL----- 468
Db 750 HTESAGVITAGAVGILLFLILVGLLIFFLKRNRKKQKQKPELRDLVFSPPGDIADFDAD 809
Qy 469 -----KQDVRSGAA-----ARGRAALLYSADDSGFELVGLASALCQLPLRVAV 514
Db 810 HVRKNERDSNCGFADEYQQLSLVGHSSQSQMASSENNAKNRYRNVLPYDWSRVPKPKTH 869
Qy 515 D-----LWSRRE--LSAQGPVAFHQRQTL-----QEGGVVLL-- 548
Db 870 EEPGSDYINASFMPCLWSPQEFIAQGPPLP-----QTVGDFWRLVWQSQSHTLVMLTN 922
Qy 549 -FSPGVALCSEWLDQGVSGGPAHG 572
Db 923 CMEAGRVCHEWYPLD--SQPCTHG 945

RESULT 9
T00093
hypothetical protein KIAA0467 -- human (fragment)
C:Species: Homo sapiens (man)

Query Match	3.08;	Score 109.5;	DB 2;	Length 2055;
Best Local Similarity	20.44;	Pred. No. 9.5;		
Matches 151;	Conservative 70;	Mismatches 233;	Indels 285;	Gaps 36;
QY	9	SLALGRSPVVLISLERLVGP	ODATHCSPGLSCLRWSDS	DLCLPGDIV-----PAPGPV-LA 62
DB	698	SLALRSQEFIIYEASGPRS-		PLDMVSSRSSDAARPVGQV 737
QY	63	PTHQLTELVLRCQKETDCD	LCLRVAVHLAVHGHWE-	PEDEKEFGGAADSEVEPRNASL 121
DB	738	DRHIQ-----	LLVHGVGAGPEITDELVR	LVLCRLDE--ATL 772
QY	122	QAQVLSFOAYPTARCVLLE	VOVPAALVQGSQSVSVYD	CFAAALGSEVRINYSYTPRY 181
DB	773	DVITVM-----	LVRCKLTADVEFTQPEGS	UPSEVLHIALPITSCRW----- 815
QY	182	EKELNHTTQQLPALPWLNV	SADGNVHLVNLVNSEBQ-	---FGLSILYMNQVQPPKPR--- 234
DB	816	-----LPALAW-----	YLRQNLLIFLHSPKYTDS	NRNHFQHPLPQGGGLP 856
QY	235	-----WHKNLTGPQ-----	ITLNTIDL--VPLCICIQW	PLE-----PDSVR----- 269
DB	857	DLDIYLYNKFGGOGTGKGV	ACITLAFVDEGGAP-LS	LALWPPSPGPPDLREEFEQL 915
QY	270	TNI--CFERED-----	PRAHONLQAAARLLRLT	LQSWLLDAPCSILPAEALCWRAPG 319
DB	916	TQVIRCPVVVDSSAONGAP	RFLRDVWEKNISIVLEEK	LRGAARQALADAIITELQLLP 975
QY	320	GDPQQLVPPLSVENVTVD	YNSSSEKIQLECLWADSL	GLPKDDVLLILLETGPDQNRSLCA 379
DB	976	ASLCTEDTPTGSLRNGSL	ETKSSAGR-----AST	FPP-----A 1008
QY	380	LEPSGCTSLPSKASTRAAR	GEYLLQDLQSQCLQIWD	-----DDLGAALWCPMKYI 432
DB	1009	PVPEGPVTPPSKAGRS-		FDWMLSKTECBLGS----- 1040
QY	433	HKRWALWVLAACLLFAAA	LSLTLKKDHAKGWLRL	LKKQDVRS-CAAAARGRAALLLYSADD 491
DB	1041	-----PKTTDDIVLDRP	EDTGRRRHKTESVTPG	AGERA-----PGSD 1079
QY	492	SGFERLVGALASALCOLP	LRVAVDLWSRRELSAQ	GPVAFHAQRROT--LOEGGVVVILF 549
DB	1080	SGAQR-----		QKRRTTQLEEGEVGTL-- 1100
QY	550	SPGNAVALCSWL-----	QDGVSGPGAHCHPDA	FRASLSCVLPDF--LOGRAPGSYVGACFD 603
DB	1101	HPVFAVQAQRWMEFMVQI	CASVRSRAHVSRLF	LPILSEFTALVTSMAGDTSVRIFE 1160
QY	604	RLHHPDAVPALFRTPV	FTFLPSQLPDFLGA	LQOPRAPRS---GR--LOER--AEQVYRA 655
DB	1161	Q--HLGSEPIFGCS-----	PGQLGSPRPAAEHRH	LLLGRNFWLQWRRTTQQAARA 1210
QY	656	LQALDSYFHPPTGP	PAPGR 674	
DB	1211	MQ-----RFEFGDGSSGR	1224	

RESULT 11
sensor histidine kinase - Deinococcus radiodurans (strain R1)
C.Species: Deinococcus radiodurans
C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C.Accession: B75376
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; De
S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zall
M.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: B75376
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-433 <WHI>
A:Cross-references: GB:AE002004; GB:AE000513; NID:96459366; PIDN:AAF11166.1; PID:9645937
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1606
A:Map position: 1
C:Superfamily: hypothetical protein HII707; sensor histidine kinase homology

Query Match 2.9%; Score 107.5; DB 2; Length 433;
Best Local Similarity 22.8%; Pred. No. 1.9;
Matches 106; Conservative 57; Mismatches 154; Indels 147; Gaps 24;

Qy 289 ARRLTLTQSWLLDAPCSLPAEALCWAPGGD-----PCQPLVPPLSWENVTVDVNS 341
Db 5 ARRLTLTQSWLLDAPCSLPAEALCWAPGGD-----PCQPLVPPLSWENVTVDVNS 341
Qy 342 SEKL-----OLOECLWADSL-GPLKDDVLLLETRGPQDNRSICALPSPGCTSL 388
Db 65 ADILEKGTAAAGRIYQGTQLRWGGADGPPVLDAAFLQTSEPR-----AVERGDLV 118
Qy 389 PSKASTRAARLGEYLQDLSGQCLQWLDDDLGALWACPMKDYTHKRWALVWLACILFAA 448
Db 119 ASR-----ROGEYV---IQVGRNLE-----PLE-HLLRRYALLSGLSLL-- 153
Qy 449 ALSLILLLKDHAKGWL---RLKQDVRSNAAGRAALL-----LYSADDSGFERLV 498
Db 154 ALSLL-----AGWVAKQVRLTLPESLVGRVQQLDSPDP 199
Qy 499 GALASALCOLPLRVAVDLWSRRLSAGQPVAFHQR-ROTLQEGGVVVLFSFGAVALC 557
Db 200 GALARAL-----OSSLDALRAERETFLASASHELRTPTVTAMLA 240
Qy 558 SEWLQDQVSGGAGHPDAPFRASLSCLVPLDFLOGRAGSVYVAGCFDRLHLPDAPAL--- 614
Db 241 D-VQHTLSRP---RPPDELRAAL-----MRAEQTASRLRLTLGNLMTLWVQSLPEP 288
Qy 615 ---FRTVPVFTLPSQLPD-----FLGALQPPAPRSGRLQERAEQ-VSRALQ 657
Db 289 RAARWOLDLHLAGEAVDRLQPLSLGRVGLWLDGQAPVSGEAGLLSSVAENLIGNAIK 348
Qy 658 PALDSVFHPG-----TPAPGRG-----VGPAGAGGAGDG 687
Db 349 -----FTPAGGEVQVRVTPLPNGGAELVVQDTGPGPPPTLTG 386

RESULT 12
A60386
Interleukin-4 receptor precursor - human
A:Alternate names: IL-4 receptor
A:Species: Homo sapiens (man)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999
C:Accession: A60386; A47603
R:Galizzi, J.P.; Zubek, C.E.; Harada, N.; Gorman, D.M.; Djossou, O.; Kastelein, R.; Banc
Int. Immunol. 2, 669-675, 1990
A:Title: Molecular cloning of a cDNA encoding the human interleukin 4 receptor.
A:Reference number: A60386; MUID:91120547
A:Accession: A60386
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-825 <GAL>
R:Idzerda, R.L.; March, C.J.; Mosley, B.; Lyman, S.D.; Vanden Bos, T.; Gimpel, S.D.; Dir
J. Exp. Med. 171, 861-873, 1990
A:Title: Human interleukin 4 receptor confers biological responsiveness and defines a nc
A:Reference number: A47603; MUID:90171849
A:Accession: A47603
A:Molecule type: mRNA

A:Residues: 1-74, 'I', 76-825 <ID>
A:Cross-references: GB:X52425; NID:g33833; PIDN:CAA36672.1; PID:g33834
C:Genetics:
A:Gene: GDB:IL4R
A:Cross-references: GDB:118823; OMIM:147781
A:Map position: 16p12.1-16p12.1
C:Superfamily: interleukin-4 receptor; cytokine receptor homology
C:Keywords: cytokine receptor; glycoprotein; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-825/Product: interleukin-4 receptor #status predicted <MAT>
F:26-232/Domain: extracellular #status predicted <EXT>
F:34-218/Domain: cytokine receptor homology <CRS>
F:233-256/Domain: transmembrane #status predicted <TM>
F:257-825/Domain: intracellular #status predicted <INT>
F:53,98,128,134,176,209/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 2.9%; Score 107; DB 1; Length 825;
Best Local Similarity 19.0%; Pred. No. 4.6;
Matches 154; Conservative 75; Mismatches 270; Indels 312; Gaps 38;

Qy 41 LWDSDILCLPGDIVP-----APGVLAPTHLQTELVLRCQKQETDCDCLRVAVHLAVH 93
Db 103 LWAGQQLLWKGSKFSEHVKPRAPGNLTVHTNVSDFLLLTWSNPYPDPNLYNLHTYAVN 162
Qy 94 GHWEPEDEEFGGAADSGVEEPRNASLQAQVVLFSQAYPTAKVLE--VOVPAALVOF 151
Db 163 -IWSNDPAD-----FRYI---NVITYLESLRIAASLTLS 193
Qy 152 GOSVGVVYDCFEAALGSEVRIMSYTOPRYEKELNHTQOLPALPWLNVASDGNVHLVNL 211
Db 194 GIS-----YRARVRAAOY-----NTWSEWSPSTKWHNSYREDFEQHLLIG 236
Qy 212 VSEQHFGLSL-----YMNQVQGPKKR-----WHKNLTGQOI 244
Db 237 VSVSIVILAVCLLCYVSTIKKEWDOI PNPASRLVAIIIDAGSGQWKEKRSRGOEP 296
Qy 245 ITLNH-----TDLPCLCIQVWPLEPDSVVRNICPFREDPR-----AHQNLWQAA 289
Db 297 ACQPHWKNCLTKLLPCF-----LEHNMKRD-----EDPHKAAKEMPFQSGKSAWCP 344
Qy 290 RLRLTLQSW-----LLDAPCSLPAEALCWAPGGDPCQPLVPPLSWENVTVD 338
Db 345 EISKTVL--WPESISVVRVRCVELFEAPVECEEEV--EEEKSFC---ASPSSRDDFQE 397
Qy 339 VNSSEKLQQLQCLWADSLGPKLDDVLLLETRG---PQDNRSICALPSPGCTSL 388
Db 398 GREGIVARLTESLFDLLG-----EENGFCQDDMGESCLLPSPGSTAHMPWDEF 448
Qy 389 PSKASTRAARLGEYLLQDL-----OSGQCLQWLDDDL-----GALWACP 427
Db 449 PSAGKEAPPWKGKQPLHLEPSPASPSTQSPDNLTCTETPLVIAGNPAYRFSFNSLSQSP 508
Qy 428 MDKYL-----HKRWALVWLACILFAAALLKLDKDHAKGWLRLKQDV-RSGAAA 478
Db 509 CPRELGPDPILLARHLEVEPEMPCV---PQLSEPTTVPOPEPETWEQILRRNVLOHGA 565
Qy 479 RGRAALLLSADDSGFERLVGALASALCOLPLRVAVDLWSRRLSAGQPVAFHQROROT 538
Db 566 AP-----VSAPTSYQFEV-----HA----- 581
Qy 539 LQEGG-----VVVLLFSPGAVA-----LCSEWLQDQVSGGAGHPDAPFRASLSCLV 587
Db 582 VEQGTQASAVVGLGPGGEGAGYKAFSSLLASSAVSPEKCGFCASSGEEGYKPF---- 636
Qy 588 FLQGRAPGSYGCADFRLHLPDAPVLPALFTVPVFT-----LPSQLPDFLG 632
Db 637 LIPG-CPGD-----PAPVP-----VPLFTFGLDREPPRSPQSSHLFSSSSPHLG 679
Qy 633 A-----LQOPRAPRS-----GRQERAEQVSRALQPAL 660
Db 680 LEPEGKEVDMKPPPLPQSOATDPLVDSLGSIGVYSALTCHLCGHLKQCHGQDGGTQPM 739

QY 661 DS-----YFHPPTPAPRGVGPAGP 682
DB 740 ASPCCGCCCDRSPPTPLRAPDPSGGVP 770

RESULT 13
A83583
probable biotin synthesis protein BioC PA0503 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: A83583
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Ras, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337
A:Accession: A83583
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-274 <STO>
A:Cross-references: GB:AE004487; GB:AE004091; NID:g9946361; PIDN:AAG03892.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0503

Query Match 2.9%; Score 106.5; DB 2; Length 274;
Best Local Similarity 27.2%; Pred. No. 1.2;
Matches 92; Conservative 30; Mismatches 111; Indels 105; Gaps 22;

QY 371 PDNRSCLALEPSGCTSLFSK-----ASTRAAR-----LGEYLLQDLQSGCQLQ 414
DB 2 PDDSSPL--LAPHGVAALPKQVAASFSRAASYDAVLELQGVGESLLSALPEGSPR 59
QY 415 LWDGDLGALWACP---MDKYIHRNALVWLACLLFAALSLILLKKDHAKGWLRLKQD 471
DB 60 RW-VDLG---CGTGYFSRALER-----FGAAEGLAV-----DIAEGMLR----- 95
QY 472 VRSGAARGAALLLSADDSFRLVGLASALCOLPLR-VAVDLMSRRLSQAQPVAV 530
DB 96 ---HARAG-----GASHFEGDAE---RLPLRDGSCDL-----LFSLSLAIQW 132
QY 531 -----FHAQRROTLEGGVWLLFSPGAVALCSEWLDGVSFGCAHGHDAFRASLSGV 584
DB 133 CADLPAVLAFAARVLRPGG---VLAFLSLCVGTIGE-LRDSWRVVDGFGVHVRFRFA----- 184
QY 585 LPDFLOGRAGSVVYGACF---DRLLH-PD--AVPALFTVVFVFTLPOLDFLGLAQOQR 638
DB 185 FADYLQHAAGSGLLPLTLRHEDRLHFPDLRLSLTHLKGALHNLNPRPDGL----- 237
QY 639 APRSGRLOERAQVSRALQPALDSYFHPGCTPAPRGV 676
DB 238 ---TGR-----QRIRALVAAYERFQPEGLPATYRVV 266

RESULT 14
B75262
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: B75262
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: B75262
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-679 <WHI>
A:Cross-references: GB:AE002083; GB:AE000513; NID:g6460360; PIDN:AAF12085.1; PID:g646036

A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2546
A:Map position: 1

Query Match 2.9%; Score 105.5; DB 2; Length 679;
Best Local Similarity 26.1%; Pred. No. 4.7;
Matches 81; Conservative 20; Mismatches 111; Indels 111; Gaps 17;

QY 438 LVMACLLFA-----AALSILLKLDHAKGWLRLKQDVRSG-----AAARGAAL 484
DB 71 LVMPLAALYGNRQLAGLLPLPQQFGVGRALLRADTRRRSPADPPFAVVRGELA 130
QY 485 LLYS-----ADDSGFERLVGALASALCOLPLRVAVDLMSRRLSQAQ- 525
DB 131 LAFAEHLAAGHHMGDVRPEHALVSDAGEVRLVGADYAL-----HLAGRD 175
QY 526 --GPV--AWFHAQRROTLEGG-----GVVLLFSPGAVALCSEWLDGVSFGP 569
DB 176 FAGPVASAEYLPPEQRVGSATGAGADAESDAFLAVLFE-----LLGRHPYAGIOARG 231
QY 570 -AHGPHDAFRASLSCLVLPDFLOGR--APGSYVGCACFDRLHDPDAVPAF--RTVPVFTL 623
DB 232 AAPGGEATAAGLFVDAPQAGFGRTAPGWPFAAL-----PPAVQALEVQAFVAVVPR 286
QY 624 PSOLPD-----FLGALQQ-----PRAPSRGRLQERAEQVSRALQPALDSYFHPGCTPAPG--- 673
DB 287 PS--PETWAALCALAAELVPCARRAG-----HWQVGLPCPCSCAA 325
QY 674 --RGVGPAG 681
DB 326 EREGAPSSG 335

RESULT 15
B49555
enhancer of split homolog R-esp2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 26-May-2000
C:Accession: B49555
R:Schmidt, C.J.; Sladek, T.E.
J. Biol. Chem. 268, 25681-25686, 1993
A:Title: A rat homolog of the Drosophila enhancer of split (groucho) locus lacking
A:Reference number: A49555; MUID:94064640
A:Accession: B49555
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-741 <SCH>
A:Cross-references: GB:L14463; NID:g294548; PIDN:AAC37640.1; PID:g294549
A:Note: authors translated the codon TCC for residue 176 as Tyr, also Ala was found
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
F:497-530/Domain: WD repeat homology <WD1>
F:583-616/Domain: WD repeat homology <WD2>
F:665-698/Domain: WD repeat homology <WD3>
F:706-739/Domain: WD repeat homology <WD4>

Query Match 2.9%; Score 105.5; DB 2; Length 741;
Best Local Similarity 19.3%; Pred. No. 5.3;
Matches 104; Conservative 54; Mismatches 147; Indels 235; Gaps 27;

QY 23 RLVGQDATHCSFGISLRLWDSILCLPGDIYPAG--PVLAP--THLOTETVLRCQKET 78
DB 297 RTDAPTGTGNSNTPGLR-----PVPKPGVDPDLASSLRTPMAVPCPYPT 340
QY 79 DCDCILRVAVHLAVHGWEEPEDEKFGGAADSGVBEPRNASLQAQVLSFQAYPTARCV 138
DB 341 P-----FGIVPHAGMNGELTSP-----CAAYAGL---HNISFQMSAAA----- 376
QY 139 LLEVQVPAALVQFGSGSVVYDCFEALGSEVRINWSYTPRYEKELNHTQOLPALP--- 195
DB 377 -----AAAAAAAYGRS-----PVGFGDPHHMRVPAIPPNL 407


```

QY 196 -----WLNYSADGDNHLVLNVESEQHFGSLTYMNVOYQPPKPRMKNLTGP---- 242
Db 408 TGIPGKPAYSFHVSADG-----OMQPYFFP--PDALIGPGIP 443
QY 243 -----QITLNTDLVPCL-----CLOVMPLEPDSVRTNICPFREDPRAHQ 283
Db 444 RHARQINTLNHGEVYCAVTISNPTRHVYTGKGCYKAMDITDPGNKSPVSQL--DCLNRD 501
QY 284 NLMQARL-----RLT-----TLOSMLDAP-----CSLPAEAAICWRAPGDDPCQPLV 327
Db 502 NYIRSCRLPDGRTLIYVGEASTISIMDLAAPTPIKAELTSSAPACY----- 549
QY 328 PPLSMENYTVVNSSEKIQLOEC-----LMADSLGPKDVLLETRGPQDNRSICALE 381
Db 550 -----ALAI SPD SKVCFSCCSDGNIAW-----DLHNOTLVROFGHTDGASCIDIS 596
QY 382 PSGCTSLPSKASTRAARLGEYILODLOSOCLOLMD--DDLGLMNCPM----- 428
Db 597 NDG-TKLMTGGLDNIVR-----SMDLREGROLQOHFTSQTSLGICPTGEMLAVGMENS 650
QY 429 -----DKY--IHKRWALWLAACLFAAALSLILLKKDHAKGWLRLKODVRSQ 475
Db 651 NVEVLHTKPKDKYQLHLHE-----SCVLSLKFAGCKWF-----VRBG 688

```

Search completed: May 19, 2002, 14:11:38
 Job time: 6406 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 13:26:40 ; Search time 89.34 Seconds
(without alignments)
855.370 Million cell updates/sec

Title: US-09-608-918-10
Perfect score: 3692
Sequence: 1 MPVPWELLSLALGRSPVLS.....TPARGVGPGAGPGAGDGT 688

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 0%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SID55/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID55/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID55/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID55/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID55/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID55/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID55/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID55/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID55/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID55/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID55/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID55/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID55/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID55/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID55/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID55/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID55/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID55/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID55/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID55/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID55/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3692	100.0	688	22	Chimeric Zcytor14
2	3673.5	99.5	705	22	Human PRO polypept
3	3673.5	99.5	705	22	Human Interleukin
4	3673.5	99.5	705	22	Human PRO20040. H
5	3673.5	99.5	705	22	Chimeric Zcytor14
6	3606.5	97.7	675	22	Chimeric Zcytor14
7	3588	97.2	692	22	Human cytokine rec
8	3075	83.3	575	22	Human variant zcyt
9	1432	38.8	309	21	Human secreted pro
10	1305	35.3	267	22	Human membrane or
11	1300	35.2	332	22	Human protein sequ

12	1079	29.2	204	22	AAE10920	Human gene 12 enco
13	305.5	8.3	617	21	AAV76048	Murine skin cell p
14	305.5	8.3	617	22	AAB55987	Skin cell protein,
15	287.5	7.8	667	22	AAU04957	Human Interleukin
16	176	4.8	866	17	AAW04185	Human Interleukin
17	176	4.8	866	19	AAW61272	Human IL-17R prote
18	176	4.8	866	20	AAW92409	Human IL-17R prote
19	176	4.8	866	21	AAV99941	Human IL-17R prote
20	176	4.8	866	21	AAV97131	Human Interleukin
21	176	4.8	866	21	AAV97181	Human Interleukin
22	176	4.8	866	21	AAB03807	Human Interleukin
23	176	4.8	866	22	AAB62066	Human IL-17R (hCTL
24	176	4.8	866	22	AAV72754	Human Interleukin
25	166	4.5	864	17	AAW04184	Murine Interleukin
26	166	4.5	864	19	AAW61271	Mouse Interleukin
27	166	4.5	864	20	AAW92408	Murine IL-17R prot
28	166	4.5	864	21	AAV99935	Murine IL-17R prot
29	166	4.5	864	21	AAV97130	Murine Interleukin
30	166	4.5	864	21	AAV97180	Murine Interleukin
31	166	4.5	864	21	AAB03806	Murine Interleukin
32	166	4.5	864	22	AAB62060	Murine IL-17R poly
33	166	4.5	864	22	AAV72748	Murine Interleukin
34	147	4.0	330	21	AAV75947	Murine skin cell s
35	147	4.0	330	22	AAB55886	Skin cell protein,
36	126.5	3.4	1711	19	AAW70506	Osteostesticular pr
37	126.5	3.4	1711	19	AAW70507	Mutant osteostetic
38	116	3.1	5435	22	AAE10145	Streptomyces nous
39	115.5	3.1	2424	22	ABBS8934	Drosophila melanog
40	113	3.1	911	22	AAU69475	Human purified sec
41	113	3.1	4924	22	AAV70968	S. spinosa protein
42	113	3.1	4928	20	AAV39300	Spnd a polyketide
43	112.5	3.0	840	21	AAV93652	A mammalian solubl
44	112.5	3.0	840	21	AAV70028	Soluble Interleuki
45	111.5	3.0	586	22	AAU17979	Human immunoglobul

ALIGNMENTS

RESULT 1
AAB61883
ID AAB61883 standard; Protein; 688 AA.
XX AAB61883;
AC
DT 08-MAY-2001 (first entry)
XX
XX Chimeric Zcytor14 protein #1.
DE
XX Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;
KW antiinflammatory; gene therapy; vaccine.
XX Homo sapiens.
OS
XX WO200104304-A1.
PN
XX 18-JAN-2001.
PD
XX 30-JUN-2000; 2000WO-US18383.
PF
XX 07-JUL-1999; 99US-0348854.
PR
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Presnell SR, Burkhead SK, Powder SL;
PI
XX WPI; 2001-112618/12.
DR
XX New polypeptide encoding a human cytokine receptor Zcytor14, for
PT treating inflammation e.g. rheumatoid arthritis -
XX Claim 2; Page 102-104; 112pp; English.
PS
XX

CC The invention provides a new human cytokine receptor designated Zcytor14.
 CC Zcytor14 can be expressed by standard recombinant methodology. The
 CC encoding nucleic acid is useful for detecting the expression of a
 CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be
 CC used to screen biological samples in vitro for the presence of Zcytor14.
 CC proteins, polypeptides and peptides having Zcytor14 activity can be
 CC administered to a subject who lacks an adequate amount of this
 CC polypeptide, for treating inflammation and conditions such as rheumatoid
 CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14
 CC antibodies) can be used to treat a subject who produces an excess of
 CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide
 CC Zcytor14 to a subject. The present sequence represents a chimeric
 CC Zcytor14 protein.

XX Sequence 688 AA;

Query Match 100.0%; Score 3692; DB 22; Length 688;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 688; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVPNFWLLSLGRSPVLSRLVGPQDATHCSPLGSLRLWSDILCLPGDIVPAPGPV 60
 DB 1 mpvpnwllslgrspvlsrlvqpqathcsplgslrldsdilclpgdivpapgpy 60
 QY 61 LAPTHLOTLVLCQKETDCDCLRVAVHLAVHGHWEPEDEKFGGAADSGVEEPRNAS 120
 DB 61 lapthlotlvlcqketcddclrvavhlavhghwepeedeekfggaadsgveeprnas 120
 QY 121 LQAQVVLFOAYPTARCVLEVOVPAALVQFGOSGVSVVYDCFAALGSGEVRWTSYQPR 180
 DB 121 lqaqvvlfoayptarcvlevoqpaaalvqfgosgvsvvydcfaalgsgvriwtsyqpr 180
 QY 181 YKELNHTQOLPALPWLNVNSADGNDVHLVNVSEEHFGLSLXWNVQGGPKPRWHKNLT 240
 DB 181 ykelnhtqolpalpwlvnvnsadgndvhlvlnvseehfglslxwnvqgppkprwhknl 240
 QY 241 GPQITLNTDLVPLCIQWPLEPDSVRTNICPFREDPRAHNLWQAARLLTLQSWL 300
 DB 241 gpqitlntdlvplciqwplepdsrvrtnicpfredprahnlwqaarlrltlqswl 300
 QY 301 LDAPCSLPAEALCWRAPGDCQPIVPLPSWENVTVDVNSSSEKLOLQECIWDADSLGPK 360
 DB 301 ldapcslpaealcwragpdcqpilvplpswenvtvdvnsssekloleciwadslgplk 360
 QY 361 DDVLLLETRGPQDNRSICALFSGCTSLPSKASTRAARLGEYLLQDLQSGCQLQWDDDL 420
 DB 361 ddvllletrgpqdnrsicalfsgctslpskastaarlgelyllqdlqsgcqlqwdddl 420
 QY 421 GAIWACPMCKYTHKRWALVWVLAACLLFAAALSLILLKDKHAKGWLRLKODVRSGAARG 480
 DB 421 galwacpmckythkrwalvwvlaacllfaaalslillkdkhagwlrllkqdvrsgaarg 480
 QY 481 RAALLYSADDSGFERLVGALASALCQLPLRVAVDLSRRELQAQPGVAVFHAQRRTLQ 540
 DB 481 raallysaddsgferlvgalasalcqlplrvavdlswrelsaqpgvavfhaqrtrtlq 540
 QY 541 EGGVVVLLFSPGAVALCSWLDQGVSGPGAHGPHDAFRASLSCVLPDFLOGRPGSVYGA 600
 DB 541 eggvvllfsgpavalcswldqgvsgpgahghpdafraaslscvlpdflogrpgsvyga 600
 QY 601 CFDRILLHPDAVPALFRTVPFTLPSPDLFLGALQOPRAPRSGRLQERAEQVSRALQPAL 660
 DB 601 cfdriillhpdavpalfrtvpftlpspdlflgalqopraprsgrlqeraeqvsralqpal 660
 QY 661 DSYFHPGTPAPRGVGPAGPGAGDGT 688
 DB 661 dsyfhpptpapgvgpapgagdgdt 688

RESULT 2

AAU29322

ID AAU29322 standard; Protein; 705 AA.

XX AC AAU29322;
 XX DT 18-DEC-2001 (first entry)
 XX DE Human PRO polypeptide sequence #299.
 XX KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200168848-A2.
 XX PD 20-SEP-2001.
 XX PF 28-FEB-2001; 2001WO-US06520.
 XX PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 03-MAR-2000; 2000US-187202P.
 PR 06-MAR-2000; 2000US-186968P.
 PR 14-MAR-2000; 2000US-189320P.
 PR 14-MAR-2000; 2000US-189328P.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 21-MAR-2000; 2000US-190828P.
 PR 21-MAR-2000; 2000US-191007P.
 PR 21-MAR-2000; 2000US-191048P.
 PR 21-MAR-2000; 2000US-191314P.
 PR 28-MAR-2000; 2000US-192655P.
 PR 29-MAR-2000; 2000US-193032P.
 PR 29-MAR-2000; 2000US-193053P.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 04-APR-2000; 2000US-194449P.
 PR 04-APR-2000; 2000US-194647P.
 PR 11-APR-2000; 2000US-195975P.
 PR 11-APR-2000; 2000US-196000P.
 PR 11-APR-2000; 2000US-196187P.
 PR 11-APR-2000; 2000US-196690P.
 PR 11-APR-2000; 2000US-196820P.
 PR 18-APR-2000; 2000US-198121P.
 PR 18-APR-2000; 2000US-198585P.
 PR 25-APR-2000; 2000US-199397P.
 PR 25-APR-2000; 2000US-199550P.
 PR 25-APR-2000; 2000US-199654P.
 PR 03-MAY-2000; 2000US-201518P.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 30-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 05-JUN-2000; 2000US-209832P.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 22-AUG-2000; 2000US-0644848.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 XX (GETH) GENENTECH INC.
 PA Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX WPI; 2001-602746/68.
 DR N-PSDB; AAS46223.
 XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 PT presence of tumours, such as prostate and breast tumours, in mammals and
 PT to screen for modulators of the compounds -
 XX Claim 11; Fig 598; 774pp; English.
 PS

RESULT	3
AAU04956	
ID	AAU04956 standard; Protein: 705 AA.
XX	
AC	AAU04956;
XX	
DT	24-OCT-2001 (first entry)
XX	
DE	Human Interleukin 17 receptor, IL-17RH2.
XX	
KW	Human; Interleukin-17 receptor; IL-17RH2; agonist; antagonist;
KW	PRO20040: DNA 164625-2890: systemic lupus erythematosus;
KW	rheumatoid arthritis; osteoarthritis; diabetes mellitus;
KW	allergic disease; asthma; demyelinating disease;
XX	degenerative cartilaginous disorder; transplantation associated disease
XX	
OS	Homo sapiens.
XX	

OS	Homo sapiens.	Location/Qualifiers
XX		1..20
FT	Key	/label=Signal-peptide
FT	Peptide	21..705
FT	Protein	/label=Mature-II-17RH2
FT		

FT		/note= "N-myristoylation site"
FT	Modified-site	118..121
FT		/note= "Asn is N-glycosylated"
FT	Region	152..157
FT		/note= "N-myristoylation site"
FT	Modified-site	186..189
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	198..201
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	211..214
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	238..241
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	248..251
FT		/note= "Asn is N-glycosylated"
FT	Region	319..324
FT		/note= "N-myristoylation site"
FT	Modified-site	334..337
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	357..360
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	391..394
FT		/note= "Asn is N-glycosylated"
FT	Region	438..443
FT		/note= "N-myristoylation site"
FT	Domain	453..473
FT		/note= "Transmembrane domain"
FT	Region	516..521
FT		/note= "N-myristoylation site"
FT	Region	552..555
FT		/note= "CAMP/GMP-dependent protein kinase phosphorylation site"
FT	Region	583..586
FT		/note= "Glycosaminoglycan attachment site"
FT	Region	612..617
FT		/note= "N-myristoylation site"
FT	Region	692..697
FT		/note= "N-myristoylation site"
FT	Region	696..701
FT		/note= "N-myristoylation site"
FT	Region	700..705
FT		/note= "N-myristoylation site"
XX		
XX		
PX		
PD		
PN		
MO	200146420-AZ.	
28-JUN-2001.		
20-DEC-2000; 2000WO-US34956.		

```

XX 23-DEC-1999; 99US-0172096.
PR 30-DEC-1999; 99WO-US31274.
PR 11-JAN-2000; 2000US-0175481.
PR 18-FEB-2000; 2000WO-US04341.
PR 02-MAR-2000; 2000US-0505841.
PR 21-MAR-2000; 2000US-0191007.
PR 02-JUN-2000; 2000WO-US07532.
PR 22-JUN-2000; 2000US-0213087.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 24-OCT-2000; 2000US-0242837.
PR 10-NOV-2000; 2000WO-US30873.
PR 28-NOV-2000; 2000US-0253646.
PR 01-DEC-2000; 2000WO-US32678.
XX
PA (GETH ) GENENTECH INC.
XX
XX Chen J, Filvaroff E, Pong S, Goddard A, Godowski PJ, Grimaldi CJ;
PI Gurney AL, Li H, Hillan KJ, Tumas D, Van Lookeren M, Vandlen RL;
PI Watanabe CK, Williams FW, Wood WI, Yansura DG;
XX
DR WPI; 2001-451708/48.
DR N-PSDB; AAS09515.
XX
PT Novel PRO polypeptides homologous to interleukin-17, useful for the
PT diagnosis and treatment of immune related disease e.g. rheumatoid
PT arthritis and diabetes -
XX
PS Claim 10; Fig 14; 18pp; English.
XX
CC The sequence is PRO20040 which is the human Interleukin 17 receptor,
CC IL-17RH2, encoded by DNA 164625-2890. A composition
CC containing ant/agonists to the PRO polypeptides or individual components
CC are useful for treating a mammal with an immune related disease, e.g.
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopaenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease, an autoimmune or immune-mediated skin
CC disease, contact dermatitis, an allergic disease e.g. food
CC hypersensitivity, asthma, a transplantation associated disease, or a
CC chronic inflammatory demyelinating polyneuropathy. Treating a
CC degenerative cartilaginous disorder comprises administering a PRO1031 or
CC PRO1122 polypeptide agonist, or antagonist to the mammal. Numerous
CC examples of the diseases and disorders are given in the specification.
XX
XX Sequence 705 AA;
SQ
Query Match 99.5%; Score 3673.5; DB 22; Length 705;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 688; Conservative 0; Mismatches 0; Indels 17; Gaps 1;
QY 1 MPVPWFLSLALGRSPVLSRLVGPQDATHCSPLGSLCRLWSDILCLPGDIVPAPGPV 60
DB 1 mpvpwfillsalgrspvlsrlvvpqdathecsplgslcrlwdsdliclpgdivpappv 60
QY 61 LAPHQLQTELRLRCQKQETCDCLRLRVAVHLAVHGHWEPEDEKFGGAADSGVEEPRNAS 120
DB 61 laphqltelrlrcqkqetcdclrlrvavhlavhghwepeedeekfggaadsgveeprnas 120
QY 121 LQAVVLSFQAVPTARCVLLEQVPAALVQFGSGVSVYDCFEALGSEVRWTSYQPR 180
DB 121 lqavvlsfqavptarcvllvqpvaalvqfgsgsvsvydcfealgservrwsytqpr 180
QY 181 YKELNHTQOLPALPWLNVNSADGDNVHLVNLVSEHGFGLSLYWNVOGPPKPRWHKNLT 240
DB 181 ykelnhtqolpalpwlvnvsadgdnvhlvlnvseehfglslynvqgppkprwhknlt 240
QY 241 GPQIITLHNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHQLMWAARLRLITLQSWL 300

```

```

DB 241 gpqiitlhtdlvpciciqwplepdsvrtnicpfredprahqnlwqaarlrlitlqswl 300
QY 301 LDAPCSLPAAALCWRAPOGDCQPLVPLSWENVTVD-----YNSSE 343
DB 301 ldapcslpaaaalcwrapogdpcqplvpplsvenvtvdkvlefpilkghpnlcvqvnsse 360
QY 344 KIQLOECLWADSLGPKDDVILLETGRGPDNRSLCALEPSGCTSLPSKASTRAARLGEYL 403
DB 361 kqlqelcswadslgplkddvillletrgpdnrsilcalepsgctslpskastraarlgelyl 420
QY 404 LQDLSGQCLQWDDDLGALWACPMDKYTHKRWALVWLACLLFAAALSLLILLKKDHAKG 463
DB 421 lqdlsggcqlqldwddlgaalwacpmdkyihkrwalvwlacllfaaalsllillkkdhakg 480
QY 464 WLRLKQDVRSAAAARGAALLLYSADDSGFERLVGALASALCQPLRVAVDLWSRRELS 523
DB 481 wlrlkqdvrsaaaargaallllysaddsgferlvgalasalcqplrvavdlwsrrels 540
QY 524 AOGPVANFHAQRQTLQEGVVVLLFSPGAVALCSEWLQDGVSGGAGHGHDAFASLSC 583
DB 541 aggpvawfhaqrqtlqegvvvllfsgpavalcsewlqdgvgvgaghhghdaftaslsc 600
QY 584 VLPDFLOGRAPGSYVACFCFDRLLHPDAVPALFPTVPVFTLPSQLPDLGALQOPRAPRS 643
DB 601 vlpdflgrapsyvgacfdrlhpdavpalfrtvpvftlpsqlpdlgalqgpraprs 660
QY 644 RLQRAEQVSRALQPALDSYFHPGTPAPGRGVGPGAGPGAGDGT 688
DB 661 rlqraeqvstralqpaldsyfhpptpapgryvgpgagpgagdgdt 705
XX
RESULT 4
AAB87606
ID AAB87606 standard; Protein; 705 AA.
AC AAB87606;
XX
DT 15-MAY-2001 (first entry)
XX Human PRO20040.
XX Human; PRO protein; mapping.
XX Homo sapiens.
XX WO2001163118-A2.
XX 08-MAR-2001.
XX
XX 24-AUG-2000; 2000WO-US23328.
XX
XX 01-SEP-1999; 99WO-US20111.
XX 15-SEP-1999; 99WO-US21090.
XX 07-DEC-1999; 99US-0169495.
XX 09-DEC-1999; 99US-0170262.
XX 11-JAN-2000; 2000US-0175481.
XX 18-FEB-2000; 2000WO-US04341.
XX 18-FEB-2000; 2000WO-US04342.
XX 22-FEB-2000; 2000WO-US04414.
XX 01-MAR-2000; 2000WO-US05601.
XX 03-MAR-2000; 2000US-0187202.
XX 25-APR-2000; 2000US-019397.
XX 22-MAY-2000; 2000WO-US14042.
XX 05-JUN-2000; 2000US-0209832.
XX (GETH ) GENENTECH INC.
XX
XX Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX WPI; 2001-183260/18.
DR N-PSDB; AAF92138.

```

XX Eighty four nucleic acids encoding PRO polypeptides, useful in
 PT molecular biology, including use as hybridization probes, and in
 PT chromosome and gene mapping.
 XX
 PS Claim 12; Fig 162; 278bp; English.
 XX

CC The present sequence is a human PRO polypeptide (secreted and
 CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
 CC anti-PRO antibodies are useful for preparation of a medicament useful in
 CC the treatment of a condition which is responsive to the PRO protein,
 CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
 CC employed as molecular weight markers for protein electrophoresis. The PRO
 CC coding sequence has applications in molecular biology, including use as
 CC hybridisation probes, and in chromosome and gene mapping.
 CC
 XX

Sequence 705 AA:

Query Match 99.5%; Score 3673.5; DB 22; Length 705;
 Best Local Similarity 97.6%; Pred. No. 0;
 Matches 688; Conservative 0; Mismatches 0; Indels 17; Gaps 1;

QY 1 MPVPWFLSLALGRSPVLSLELVGPQDATHCSGSLSCRLMSDITLCPGDIVAPGPV 60
 DB 1 mpvpwflslalgrspvlslelvgpqdat hcs gsl scl rlm sdi tcl p gdi v ap gpv 60
 QY 61 LAPTHLQTELVLRCKETDCDCLRAVAHLAVHGWEEPEDEKFGGAADSGVEEPNRS 120
 DB 61 lapthqltelv lr cket dcd cl rva hlvah gw ee p ee k fg gaa dsg v ee p nrs 120
 QY 121 LQAVVLSFQAYPTARCVLLEVQVPALVQFGSGSVVYDCFEALGSEVRIMSTQPR 180
 DB 121 lqavvlsfqayptarcvll evqvpalvqfgsgsvvydcfeal gsevrims tqpr 180
 QY 181 YEKELNHTQOLPALPMLNVSADGDNVHLVNVSEOHFGSLVYNOVGPPKPRHKNLT 240
 DB 181 yekelnhtqolpalpmlnvsadgdnvhlv nvseohfgslv ynvgppkpr hknlt 240
 QY 241 GPQITLNLNHTDLPVLCICIQWPLEPDSVRNTRNICPFREDPRAHONLMQAAARLLTLQSWL 300
 DB 241 gpqitlnhtdlvp lcic iqw p l e p d s v r n t r n i c p f r e d p r a h o n l m q a a r l l t l q s w l 300
 QY 301 LQAPCSLPAPALCMWARGDPCQPLVPPLSMENVTVD-----VNSSE 343
 DB 301 lqapcslpapa lcmwargdpcqplvpplsmenvtvd-----vnsse 343
 QY 344 KIQGLQECIMADSGPLKDDVLLLETRGPDNRSLCALEPSCGTSLPKASRAARLGEYL 403
 DB 344 k i q g l q e c i m a d s g p l k d d v l l e t r g p d n r s l c a l e p s c g t s l p k a s r a a r l g e y l 403
 QY 404 LQDLQSGOCLQWLMDDDLGAALMACPMKDYIHKRMALVWLACLLFAALSLILLKKDHAKG 463
 DB 404 l q d l q s g o c l q w l m d d l g a a l m a c p m k d y i h k r m a l v w l a c l l f a a l s l i l l k k d h a k g 463
 QY 421 LQGLGSGCQLQWDDDLGALWACPMKDYIHKRMALVWLACLLFAALSLILLKKDHAKG 480
 DB 421 l q g l g s g c q l q w d d l g a l w a c p m k d y i h k r m a l v w l a c l l f a a l s l i l l k k d h a k g 480
 QY 464 WLRLLKQDVNSGAARARALLLXSADSGFERLVGALASLQCLPLFVAVDMSRRRLS 523
 DB 464 w l r l l k q d v n s g a a r a r a l l x s a d s g f e r l v g a l a s l q c l p l f v a v d m s r r r l s 523
 QY 481 WTLILKQDVNSGAARARALLLXSADSGFERLVGALASLQCLPLFVAVDMSRRRLS 540
 DB 481 w t l i l k q d v n s g a a r a r a l l x s a d s g f e r l v g a l a s l q c l p l f v a v d m s r r r l s 540
 QY 524 AGCPVAMFHHOROTLOEGGVVLLFSPGVALCSEMIQDGVSGGAGHPDAPFASISC 583
 DB 524 a g c p v a m f h h o r o t l o e g g v v l l f s p g v a l c s e m i q d g v s g g a g h p d a p f a s i s c 583
 QY 541 AGGPVAMFHHOROTLOEGGVVLLFSPGVALCSEMIQDGVSGGAGHPDAPFASISC 600
 DB 541 a g g p v a m f h h o r o t l o e g g v v l l f s p g v a l c s e m i q d g v s g g a g h p d a p f a s i s c 600
 QY 584 VLDPFLQGRAGSVGACFRLRLHPDAVPALFRTVPVFTLSQLPDEFGALQOFPAPRPSG 643
 DB 584 v l d p f l q g r a g s v g a c f r l r l h p d a v p a l f r t v p v f t l s q l p d e f g a l q o f p a p r p s g 643
 QY 601 VLPDFLQGRAGSVGACFRLRLHPDAVPALFRTVPVFTLSQLPDEFGALQOFPAPRPSG 660
 DB 601 v l p d f l q g r a g s v g a c f r l r l h p d a v p a l f r t v p v f t l s q l p d e f g a l q o f p a p r p s g 660
 QY 644 RLQERAEQVSRAALQPALDSYFHPPTAPACRGVSGAGPGAGDGT 668
 DB 644 r l q e r a e q v s r a a l q p a l d s y f h p p t a p a c r g v s g a g p g a g d g t 668
 QY 661 RLQERAEQVSRAALQPALDSYFHPPTAPACRGVSGAGPGAGDGT 705
 DB 661 r l q e r a e q v s r a a l q p a l d s y f h p p t a p a c r g v s g a g p g a g d g t 705

RESULT 5

AAB61884
 ID AAB61884 standard; Protein; 705 AA.
 XX
 AC AAB61884;
 XX
 XX
 DT 08-MAY-2001 (first entry)
 XX
 XX

DE Chimeric Zcytor14 protein #2.
 KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;
 KW antiinflammatory; gene therapy; vaccine.
 XX
 OS Homo sapiens.

WO200104304-A1.

18-JAN-2001.

30-JUN-2000; 2000WO-US18383.

07-JUL-1999; 9905-0348854.

(ZYMO) ZYMOGENETICS INC.

Presnell SR, Burkhead SK, Powderer SL.

WPI; 2001-112618/12.

New polypeptide encoding a human cytokine receptor zcytor14, for
 treating inflammation e.g. rheumatoid arthritis -

Claim 2; Page 105-107; 112pp; English.

The invention provides a new human cytokine receptor designated zcytor14.
 zcytor14 can be expressed by standard recombinant methodology. The
 encoding nucleic acid is useful for detecting the expression of a
 zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be
 used to screen biological samples in vitro for the presence of zcytor14.
 CC proteins, polypeptides and peptides having zcytor14 activity can be
 CC administered to a subject who lacks an adequate amount of this
 CC polypeptide, for treating inflammation and conditions such as rheumatoid
 CC arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14
 CC antibodies) can be used to treat a subject who produces an excess of
 CC zcytor14. Zcytor14 nucleotide sequences can also be used to provide
 CC zcytor14 to a subject. The present sequence represents a chimeric
 CC zcytor14 protein.
 XX

Sequence 705 AA:

Query Match 99.5%; Score 3673.5; DB 22; Length 705;
 Best Local Similarity 97.6%; Pred. No. 0;
 Matches 688; Conservative 0; Mismatches 0; Indels 17; Gaps 1;

QY 1 MPVPWFLSLALGRSPVLSLELVGPQDATHCSGSLSCRLMSDITLCPGDIVAPGPV 60
 DB 1 mpvpwflslalgrspvlslelvgpqdat hcs gsl scl rlm sdi tcl p gdi v ap gpv 60
 QY 61 LAPTHLQTELVLRCKETDCDCLRAVAHLAVHGWEEPEDEKFGGAADSGVEEPNRS 120
 DB 61 lapthqltelv lr cket dcd cl rva hlvah gw ee p ee k fg gaa dsg v ee p nrs 120
 QY 121 LQAVVLSFQAYPTARCVLLEVQVPALVQFGSGSVVYDCFEALGSEVRIMSTQPR 180
 DB 121 lqavvlsfqayptarcvll evqvpalvqfgsgsvvydcfeal gsevrims tqpr 180
 QY 181 YEKELNHTQOLPALPMLNVSADGDNVHLVNVSEOHFGSLVYNOVGPPKPRHKNLT 240
 DB 181 yekelnhtqolpalpmlnvsadgdnvhlv nvseohfgslv ynvgppkpr hknlt 240
 QY 241 GPQITLNLNHTDLPVLCICIQWPLEPDSVRNTRNICPFREDPRAHONLMQAAARLLTLQSWL 300
 DB 241 gpqitlnhtdlvp lcic iqw p l e p d s v r n t r n i c p f r e d p r a h o n l m q a a r l l t l q s w l 300

QY 301 LDAPCSLPAAALCWRAPGDPQCPPLVPPLSWENVTVD-----VNSSE 343
DB 301 ldapcslpaaalcwragpdpqcpplvpplswenvtdkvlefllkghpnlcvqnsse 360
QY 344 KIQLOECLWADSLGPKDDVILLLETRGPQDNRSICALPEPSCGCTSLPSKASTRAARLGEYL 403
DB 361 kqlqecclwadsigplkddvillletrgpqdnrsicalpepscgctslpskastaaralgeyl 420
QY 404 LQDLSQOCLQDWDLDLWALWACPMCKY IHKRWALVWLACLLFAAALSILLKKDHAKG 463
DB 421 lqdlsgqclqldddlgallwacpmckyihkrwalvwlacllfaaalsillkkdhakg 480
QY 464 WRLLLKQDVRSGAARGAALLYSADDSGFERLVGALASALCOLPLRVAVDLWSRRELS 523
DB 481 wrlllkqdvrsagaargaalysadsgferlvgalasalcolplrvavdlwsrrels 540
QY 524 AQGPVAFHQAORQTLOEGVVVLLFSPGAVLCEWLQDGVSGPAGHPHDAFRASLSC 583
DB 541 aggpvafhaqrqtloeggvvllfsgpavalcsewlqdgvsppgaghpghdafraslsc 600
QY 584 VLPDFLQGRAPGSVVGACFDRLLHPDAVPALFRTPVFTLPSQLPDFLQALQOPRPSG 643
DB 601 vlpdfllqgrapgsyvgacfdrlhpdavpalfrtpvftlpsqlpdfllqalqopprpsg 660
QY 644 RIQERAEQVSRALQPALDSYFHPGTPAPGRGVGPGAGPGAGDGT 688
DB 661 riqeraeqvsralqpaldsyfhpptpapgrrvgpgagpgagdgdt 705

RESULT 6
AAB61885
ID AAB61885 standard; Protein; 675 AA.
XX AAB61885;
AC AAB61885;
XX
DT 08-MAY-2001 (first entry)
DE Chimeric Zcytor14 protein #3.
XX
KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;
KW antinflammatory; gene therapy; vaccine.
OS Homo sapiens.
XX
PN WO200104304-A1.
XX
PD 18-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18383.
XX
PR 07-JUL-1999; 99US-0348854.
XX (ZYMO) ZYMOGENETICS INC.
PI Presnell SR, Burkhead SK, Powder SL;
XX
DR WPI; 2001-112618/12.
XX
PT New polypeptide encoding a human cytokine receptor Zcytor14, for
PT treating inflammation e.g. rheumatoid arthritis.
PS
PS Claim 2; Page 107-109; 112pp; English.
XX
CC The invention provides a new human cytokine receptor designated Zcytor14.
CC Zcytor14 can be expressed by standard recombinant methodology. The
CC encoding nucleic acid is useful for detecting the expression of a
CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be
CC used to screen biological samples in vitro for the presence of Zcytor14.
CC Proteins, polypeptides and peptides having Zcytor14 activity can be
CC administered to a subject who lacks an adequate amount of this
CC polypeptide, for treating inflammation and conditions such as rheumatoid
CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14

antibodies) can be used to treat a subject who produces an excess of
CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide
CC Zcytor14 to a subject. The present sequence represents a chimeric
CC Zcytor14 protein.
XX
SQ Sequence 675 AA;
Query Match 97.7%; Score 3606.5; DB 22; Length 675;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 675; Conservative 0; Mismatches 0; Indels 13; Gaps 1;
QY 1 MPVPWFLLSALGRSPWLSLERLVGPQDATHSCPSGLSCLWSDILCLPGDIVPAGPV 60
DB 1 mpvpwfllsalgrspwlslerlvqpqdatcspglscrlwdsilclpgdivpapgvp 60
QY 61 LAPTHLQTELVLRQKETDCDCLRVAVHLAVHGHWEPEDEKFGGAADSGVEEPRNAS 120
DB 61 lapthlqtelvlrqketdcdclrvavhlavhghweepedeekfggaadsgveeprnas 120
QY 121 LQAQVVLSPQAYPTARCVLLEVOVPAALVQFGOSVGVVYDCFEAALGSEVRITWSYTOPR 180
DB 121 lqaqvvlspqayptarcvlllevqpaaalvqfgosvsvvydcfeaalgsevrilwscyqpr 180
QY 181 YKELNHTQQLPALPWLNVSDGDNVHLVNLVSEEHFGLSLYMNVOVGPKPRWHKNLT 240
DB 181 ykelnhtqqlpalpwlrvnsadgnvhlvlnvseeqhfglislymnvqvgpkprwhknlt 240
QY 241 GPQITLNTHTDLVPCICIQWPLEPDSVTRNICPFREDPRAHONLWQAARLLTQTQSWL 300
DB 241 gpqitlnthtdlvpcicqvwplepdsvtrnicpfredprahqnlwqaarlrltltqswl 300
QY 301 LDAPCSLPAAALCWRAPGDPQCPPLVPPLSWENVTVDVNSSEKIQLOECLWADSLGPLK 360
DB 301 ldapcslpaaalcwragpdpqcpplvpplswenvtdvnsseklqleclwadsigplk 360
QY 361 DVVLLLETRGPQDNRSICALPEPSCGCTSLPSKASTRAARLGEYLQDLSQCCLQDWDLDL 420
DB 361 dvvllletrgpqdnrsicalpepscgctslpskastaaralgeylqdlsgqclqldwdldl 420
QY 421 GALWACPMCKY IHKRWALVWLACLLFAAALSILLKKDHAKWLRLLKODVRSAGAAARG 480
DB 421 galwacpmckyihkrwalvwlacllfaaalsillkkdhak-----aaarg 467
QY 481 RAALLYSADDSGFERLVGALASALCOLPLRVAVDLWSRRELSAQGPVAFHQAORQTLO 540
DB 481 raallysadsgferlvgalasalcolplrvavdlwsrrelsaqgpvafhaqrqtlo 527
QY 541 EGGVVVLLFSPGAVLCEWLQDGVSGPAGHPHDAFRASLSCVLPDFLQGRAPGSYVGA 600
DB 541 eggvvvllfsgpavalcsewlqdgvsppgaghpghdafraslscvlpdfllqgrapgsyvga 587
QY 601 CFDRLLHPDAVPALFRTVPFTLPSQLPDFLQALQOPRPSRGLQERAEQVSRALQPAL 660
DB 601 cfdrllhpdavpalftrtpvftlpsqlpdfllqalqopprpsrgrlqeraeqvsralqpal 647
QY 661 DSYFHPGTPAPGRGVGPGAGPGAGDGT 688
DB 661 dsyfhpptpapgrrvgpgagpgagdgdt 675

RESULT 7
AAB61880
ID AAB61880 standard; Protein; 692 AA.
XX AAB61880;
AC AAB61880;
DT 08-MAY-2001 (first entry)
XX Human cytokine receptor Zcytor14.
DE Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;
KW antinflammatory; gene therapy; vaccine.

XX OS Homo sapiens.
XX PN WO200104304-A1.
XX PD 18-JAN-2001.
XX PF 30-JUN-2000; 2000WO-US18383.
XX PR 07-JUL-1999; 99US-0348854.
XX PA (ZYMO) ZYMOGENETICS INC.
XX PI Presnell SR, Burkhead SK, Pownder SL;
XX DR WPI; 2001-112618/12.
XX DR N-PSDB; AAC85027.
XX PT New polypeptide encoding a human cytokine receptor Zcytor14, for
XX PT treating inflammation e.g. rheumatoid arthritis -
XX PS Claim 2; Page 2; 112pp; English.
XX CC The invention provides a new human cytokine receptor designated Zcytor14.
XX CC Zcytor14 can be expressed by standard recombinant methodology. The
XX CC encoding nucleic acid is useful for detecting the expression of a
XX CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be
XX CC used to screen biological samples in vitro for the presence of Zcytor14.
XX CC proteins, polypeptides and peptides having Zcytor14 activity can be
XX CC administered to a subject who lacks an adequate amount of this
XX CC polypeptide, for treating inflammation and conditions such as rheumatoid
XX CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14
XX CC antibodies) can be used to treat a subject who produces an excess of
XX CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide
XX CC Zcytor14 to a subject. The present sequence represents the human
XX CC cytokine receptor Zcytor14.
XX SQ Sequence 692 AA;

Query Match 97.2%; Score 3588; DB 22; Length 692;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 675; Conservative 0; Mismatches 0; Indels 30; Gaps 2;

Qy 1 MPVPWFLSLALGRSPVLSRLERLVGPQDATHCSPLGSLRSLWSDILCLPGDIVPAPGPV 60
Db 1 mpvpwflslalgrspvlsrlervpqpqathcsplgslrslwdsdildclpgdivpappv 60
Qy 61 LAPTHLOTELRLRCOKETDCDCLRLVAVHLAVHGHWEPEDEKFGGAADSGVEEPRNAS 120
Db 61 laphlqltelvrlrcoketdcdclrlvavhlavhghweepedeekfggaadsgveeprnas 120
Qy 121 LQAOVVLSPFOAYPTARCVLLEQVPAALVQFGSGSVVYDCFEAALGSEVRWISYTPQR 180
Db 121 lqaovvlsfqaayptarcvllvqvpaaalvqfgsgsvvydcfeaalgsevrwisytcqr 180
Qy 181 YKELNHTQQLPALPWLNVYADGDNVHLVNVSEEHFGLSLXWNOVQGPKPWRHKNLT 240
Db 181 ykelnhtqqlpalpwnvsadgdnvhlvlnvseeqhfglslxwnvqgpkprwhknlt 240
Qy 241 GPQITLNTHTDVLPCIQVPLEPSVTRNICPFREDPRAHNLMOAARLLTLOSWL 300
Db 241 gpqiitlnhtdvlpciqvplepsvtrnicpfredprahnlwqaarlrltlqswl 300
Qy 301 LDAPCSLPAEALCWAPGDDPCQPLVPPLSWENVTVD-----VNSSE 343
Db 301 ldapcslpaeaalcwapggdpcqplvpplswenvtvdkvlefpllkghpnlcvqvnasse 360
Qy 344 KLQLECLWADSLGPLKDDVLLLETRGPQDNRSICALPEPGCTSLPSKASTRAARLGEYL 403
Db 361 klqlqecclwadsigplkddvlllletrgpqdnrsicalpepgctslpskastaarlgeyl 420
Qy 404 LQDLQSGQCQLQWDDDLGALWACPMDDKYIHKRWALVWLACLLFAAALLSLILLKKDHAKG 463

Db 421 lqdlsgsqcqlqwdddlgalwacpmdkyihkrwalwlacllfaaalsillllkkdhak- 479
Qy 464 WLRLKQDVRSGAAGRAALLLYSADDSGFERLVGALASALCQPLRVAVDLSRRRELS 523
Db 480 -----aaagraallllysadsgferlvgalasalcqplrvavdlwrrrels 527
Qy 524 AOGPVAFHAQRQTLOEGGVVLLFSPCAVALCSEWLQDVGSGCAHGFDAFRASLSC 583
Db 528 aqgpvawfhaqrqtloeggvvllfspgavalcsewldqvgsgpahghpdafraslsc 587
Qy 584 VLPDFLOGRAPGSYVGACFDRLHDPALFRTVPVFTLPOLPDFLGCALQOPRAPRSG 643
Db 588 vlpdflograpgsyvgacfdrlhdpavpalfrtvpvftlpqldfalgcalqopraprsg 647
Qy 644 RLQERAEQVSRLQPALDSYFHPPTGTPAPGRGVGPGAGGAGDGT 688
Db 648 rlqeraeqvsralqpaldsyfhpptgtpapgrgvpgagpgagdggt 692

RESULT 8
AAB61881
ID AAB61881 standard; Protein; 575 AA.
XX AC AAB61881;
XX DT 08-MAY-2001 (first entry)
XX DE Human variant Zcytor14 protein Zcytor14-1.
XX KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;
XX KW antiinflammatory; gene therapy; vaccine; variant; Zcytor14-1.
XX OS Homo sapiens.
XX PN WO200104304-A1.
XX PD 18-JAN-2001.
XX PF 30-JUN-2000; 2000WO-US18383.
XX PR 07-JUL-1999; 99US-0348854.
XX PA (ZYMO) ZYMOGENETICS INC.
XX PI Presnell SR, Burkhead SK, Pownder SL;
XX DR WPI; 2001-112618/12.
XX DR N-PSDB; AAC85029.
XX PT New polypeptide encoding a human cytokine receptor Zcytor14, for
XX PT treating inflammation e.g. rheumatoid arthritis -
XX PS Disclosure; Page 2-3; 112pp; English.
XX CC The invention provides a new human cytokine receptor designated Zcytor14.
XX CC Zcytor14 can be expressed by standard recombinant methodology. The
XX CC encoding nucleic acid is useful for detecting the expression of a
XX CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be
XX CC used to screen biological samples in vitro for the presence of Zcytor14.
XX CC proteins, polypeptides and peptides having Zcytor14 activity can be
XX CC administered to a subject who lacks an adequate amount of this
XX CC polypeptide, for treating inflammation and conditions such as rheumatoid
XX CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14
XX CC antibodies) can be used to treat a subject who produces an excess of
XX CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide
XX CC Zcytor14 to a subject. The present sequence represents a variant of
XX CC the human cytokine receptor Zcytor14, designated Zcytor14-1. This
XX CC variant is a truncated form of the receptor polypeptide and lacks
XX CC amino acid residues 1-113 of Zcytor14.
XX SQ Sequence 575 AA;

Query Match		83.3%;	Score 3075;	DB 22;	Length 575;
Best Local Similarity		100.0%;	Pred. No. 1.1e-274;		
Matches 575;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	114	EEPRNASLQAQVVLSTQAYTARCVLLEVOVPAALVFGVSGVSVYDCEAALGSEVRI	173		
Db	1	eeprnaslqaqvvlstqaytarcvllvqvpaaalvfqgsvgydcfeaalgservri	60		
QY	174	WSYTPRYEKELNHTQOLPALPLNVSADGDNVHLVNLVSEEQHFGLSLYWNOVGPKP	233		
Db	61	wsytpryekelnhtqolpalplnvsadgdnvhlvnlvseeqhfghslywnvqvgppkp	120		
QY	234	RHKNLTGPQIITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRHAQNLOWAARLRL	293		
Db	121	rwkhnltgpqiitlnhtdlvpcicqvwplepdsrvtnicpfredprahnqlwqaarlrl	180		
QY	294	LTLQSWLLDAPCSLPAEALCWAPAGDPCQPLVPPLSWENVTVDVNSSEKLLQECUWA	353		
Db	181	ltlqswlldapcslpaealcwragpdpqplvpplswenvtvdsnsseklqieclwa	240		
QY	354	DSLGLPKDVLLETGPDNRSLCALEPSGCTSLPSKASTRAARLGEYLLQDLSGOCL	413		
Db	241	dslgplkddvllletrgpqnrslicalepsgctslpskastraarlgeyllqdlsgqcl	300		
QY	414	QLWDDDLGALWACPMQYIHKRWALVWLACLLFAAALSILILLKKDHAKGWLRLKQDVR	473		
Db	301	qlwdddlgalwacpmdkyihkrwalvwlacllfaaalsilillkkdhakgwrlrikqdv	360		
QY	474	SGAAAGRAALLYSDDSFERLWALGALCALCOLPLRVAVDLMSRELSSAQGPVAFHA	533		
Db	361	sgeaargraallysddsferlvgalasalcolplrvavdlmsrelsaqgpvawfha	420		
QY	534	QRRTTQEGGVVLLSPGAVALCSEWLQDVGSGPAGHPDAPRASLSCVLPDFLOGRA	593		
Db	421	qrtrtqeggvvvllspgavalcsewlqdvsgpgahghpdafraslscvlpdflqgra	480		
QY	594	PGSYGACFDRLHDPDAVPAFLRTVPVFTLPSQLPDFLQALQOPRAPRSGRLQRAEQVS	653		
Db	481	pgsyvgacfdrlhdpdavnlpftrvpvftlpsqlpdfalqalqpraprsgrlqeraeqvs	540		
QY	654	RALQPALDSYFHPGTPAPGRGVPAGGAGDGT 688			
Db	541	ralqpaldsyfhpptpapgvgpaggagdgdt 575			
RESULT 9					
AA76143					
ID	AA76143 standard; Protein: 309 AA.				
XX	AA76143;				
AC					
XX					
DT	23-MAR-2000 (first entry)				
XX					
Human secreted protein encoded by gene 20.					
DE					
XX					
KW	Human; secreted protein; cancer; tumour; developmental abnormality;				
KW	foetal deficiency; blood disorder; immune system disorder; inflammation;				
KW	autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;				
KW	schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;				
KW	atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;				
KW	digestive disorder; endocrine disorder; infection; AIDS; leukaemia;				
KW	therapy; chromosome 3.				
XX					
OS	Homo sapiens.				
XX					
PN	WO9958660-A1.				
XX					
PD	18-NOV-1999.				
XX					
PF	06-MAY-1999; 99WO-US09847.				
XX					
PR	12-MAY-1998; 98US-0085093.				
XX					

PR	12-MAY-1998;	98US-0085094.
PR	12-MAY-1998;	98US-0085105.
PR	12-MAY-1998;	98US-0085180.
PR	18-MAY-1998;	98US-0085906.
PR	18-MAY-1998;	98US-0085920.
PR	18-MAY-1998;	98US-0085921.
PR	18-MAY-1998;	98US-0085922.
PR	18-MAY-1998;	98US-0085923.
PR	18-MAY-1998;	98US-0085924.
PR	18-MAY-1998;	98US-0085928.
PR	18-MAY-1998;	98US-0085925.
PR	18-MAY-1998;	98US-0085927.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;	
PI	Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;	
PI	Lafleur DW, Endress GA, Ebner R;	
XX	WPI; 2000-062296/05.	
DR	N-PSDB; AAZ65269.	
XX	New isolated human genes and the secreted polypeptides they encode,	
XX	useful for diagnosis and treatment of e.g. cancers, neurological	
PT	disorders, immune diseases, inflammation or blood disorders	
PT	Claim 11; Page 372-373; 475pp; English.	
XX	AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.	
CC	AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.	
CC	The gene encoding this protein was found to be on chromosome 3.	
CC	The genes and their corresponding secreted polypeptides are	
CC	useful for preventing, treating or ameliorating medical conditions,	
CC	e.g. by protein or gene therapy. Also pathological conditions can be	
CC	diagnosed by determining the amount of the new polypeptides in a sample	
CC	or by determining the presence of mutations in the new genes. Specific	
CC	uses are described for each of the 97 genes, based on which tissues they	
CC	are most highly expressed in, and include developing products for the	
CC	diagnosis or treatment of cancer, tumours, developmental abnormalities	
CC	and foetal deficiencies, blood disorders, diseases of the immune system,	
CC	autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive	
CC	disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin	
CC	disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney	
CC	disorders, digestive/endocrine disorders, infections and AIDS. The	
CC	polypeptides are also useful for identifying their binding partners.	
CC	The sequences shown in AAY76224 to AAY76424 represent fragments of the	
CC	secreted proteins.	
XX	Sequence 309 AA;	
SQ		

Query Match		38.8%;	Score 1432;	DB 21;	Length 309;
Best Local Similarity		90.9%;	Pred. No. 1.8e-123;		
Matches 270;		Conservative 0;	Mismatches 11;	Indels 16;	Gaps 2;
QY	1	MPVPWFLLSALGRSPVLSLRLVGPQDATHCPSGLSCLRLWSDILCLPGDIVPAGPV	60		
Db	1	mpvpwflslalgrspvlsrlrvpqpqathcpsrlscrlwdsdlclpgdivpaggpv	60		
QY	61	LAPTHLQTELVLRCQKETDCDCLRLRVAVHLVHGHWEPEDEKFGGAADSGVEEPRNAS	120		
Db	61	laphthlqtelvlrcqketdcdclrvxvhlavhghwepeedeekfggaadlgsveeprnas	120		
QY	121	LQAQVLSFQAYPTARCVLLEVOVPAALVQVGSGVSYVDVCFEALGSEVRIWSTQPR	180		
Db	121	lqaqvlsfayptarcvllvqvpaalvqvgsgvvydcfealgservrlwstqpr	180		
QY	181	YEKELNHTQOLP-----ALPWLNVSDAGDNVHLVNLVSEEQHFGLSLYWN	225		
Db	181	yekelnhtqqlpdcrglevnswpsicwlpwlvnvsadgdnvhlvnlvseeqhfghslywn	240		
QY	226	QVQGGPKPRWHKNTLGPQIITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAH	282		

Db 241 gvgppkprwhkntgqjllnhtdlvpcldqvwplepdsrvrtsapsgrtp-ah 296

RESULT 10

AA888448

ID AAB88448

AC AAB88448

DT 23-MAY-2001 (first entry)

DE Human membrane or secretory protein clone PSEC0233.

DE Human; secretory protein; membrane protein; vaccine; gene therapy;

KW rheumatoid arthritis; diabetes.

OS Homo sapiens.

PN EP1067182-A2.

PD 10-JAN-2001.

PF 07-JUL-2000; 2000EP-0114090.

PR 08-JUL-1999; 99JP-0194179.

PR 11-JAN-2000; 2000JP-0118775.

PR 02-MAY-2000; 2000JP-0183766.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;

DR WPI; 2001-093989/11.

DR N-PSDB; AAF93875.

PT Nucleic acids encoding secretory proteins/membrane proteins, useful in

gene therapy or as candidate target molecules in drug development -

Claim 1; SEQ ID 264; 609bp + CD ROM; English.

CC This invention relates to nucleic acid sequences AAF93744 - AAF93916

CC which encode human secretory or membrane proteins represented by

CC AAB88317 - AAB88419. Included in the invention are primers

CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the

CC cDNA sequences of the invention. The invention also includes methods for

CC the production of antibodies directed against the proteins, and cDNA

CC sequences, which can be used in vaccines. The polynucleotide sequences

CC can be used in gene therapy. The polynucleotide sequences and the

CC proteins they encode may be used in the prevention, treatment and

CC diagnosis of diseases associated with inappropriate secretory

CC protein/membrane protein expression. The nucleic acids and complementary

CC sequences may also be used as DNA probes in diagnostic assays

CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the

CC presence of similar nucleic acid sequences in samples. They may also be

CC used to study the expression and function of secretory proteins/membrane

CC polypeptides and their role in metabolism. The polypeptides may be used

CC as antigens in the production of antibodies against them and in assays to

CC identify modulators (agonists and antagonists) of expression and

CC activity. The antibodies and antagonists may also be used as therapeutic

CC agents to down regulate expression and activity. The antibodies may also

CC be used as diagnostic agents for detecting the presence of the

CC (EUSA). Examples of diseases which may be treated include rheumatoid

CC arthritis and diabetes.

CC Sequence 267 AA:

XX

SO

Query Match 35.3%; Score 1305; DB 22; Length 267;

Best Local Similarity 89.2%; Pred. No. 7.6e-112;

Matches 249; Conservative 1; Mismatches 17; Indels 12; Gaps 2;

OY 1 MPVPWFLSLAIGSPFVLSLERTLVGPDAITHCSPGISCRIMSDILCLPGDIVAPAGPV 60

Db 1 mpvpfllslalgrepvnslertlvpgqdatncpsplscrlwssdlclpgdtvppagpv 60

OY 61 IAPTHLQTELVLRCQKETDCDCLRLVAVHLAVHGWEEPEDEKFCGAADSGVEEPRNAS 120

Db 61 Iapthqltelvlrcqketdcclrlvavhavlavghweepedeekfggaadsgveeprnas 120

OY 121 LQAOVVSFGAATPTARCIVLLEVOVPALVQFGSGVSVYYDCFEALSGSEVRWSTQTR 180

Db 121 lqagvvlstfgayptarcvllvqvpaalvqfgsgvsvvyddcfeaalsgsevrwstqtr 180

OY 181 YEKELNHTQQLPALPMLANVSADGDNVHLVLANSSEOHFGSLTYWNOGPPKPRMKNTL 240

Db 181 yekelnhtqqlpalpmlvnvsadgdnvhlvlnvseeqhfgslstlywnvgvqppkprwhknlv 240

OY 241 GPQITLNTHTDLVPCLCIQWPLEPDSVRTNICPFREDP 279

Db 241 rpp-----psqvhshc-----rppmrvgrtqhclredp 267

RESULT 11

AA25864

ID AAM25864

AC AAM25864

DT 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:1379.

DE Human; cancer; ulcer; HIV infection; human immunodeficiency virus;

DE anti-inflammatory; antirheumatic; antiarthritic; immunosuppressive;

DE antibacterial; endocrine; cardiant; central nervous system; viricide;

DE anti-HIV; fungicide; antitumor; cardiovascular; antianemic; anaemia;

DE antiagregant; haemostatic; vulnerrary; antilucer; osteopathic; eczema;

DE dermatological; antiallergic; antidiabetic; antidiabetic; cytostatic;

DE neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;

DE immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;

DE antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;

DE cardiac dysfunction; neuropathologic disorder; cardiac anaphylaxis; autoimmune;

DE genetic disease; haematopoietic disorder; platelet disorder; asthma;

DE thrombocytopaenia; osteoporosis; severe combined immunodeficiency;

DE allergic rhinitis; diabetes; multiple sclerosis; depression;

DE Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;

DE neurological disorder.

OS Homo sapiens.

PN WO200153455-A2.

PD 26-JUL-2001.

PF 22-DEC-2000; 2000WO-US35017.

PR 23-DEC-1999; 99US-0471275.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

XX (HSE-) HSESEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457603/49.

XX N-PSDB; AAH99805.

XX Isolated human polynucleotides encoding polypeptides, useful for the

XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

XX Claim 20; Page 283; 1217pp; English.

XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to

XX AAM25963. The proteins can have activities based on the tissues and

XX cells they are expressed in, such as: antineoplastic; antirheumatic;

CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
 CC central nervous system; viricide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antinaemic; antiagregant; haemostatic; vulnerary;
 CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 XX
 SQ Sequence 332 AA;

Query Match 35.2%; Score 1300; DB 22; Length 332;
 Best Local Similarity 86.4%; Pred. No. 3e-111;
 Matches 248; Conservative 5; Mismatches 12; Indels 22; Gaps 3;
 QY 1 MPVPFLLSLALGRSPVLSLERLVGPQDATHGSPGLSCRLWDSIDLCLPGDIVPAPGPV 60
 DB 2 mpvpfllslalgrspvlslerlvqpqdhcspglscrlwdsidilcplgdvrapgpv 61
 QY 61 LAPTHLQTELVLRCQKETDCDLCLRVAVHLAVHGWEEPEDEKFGGAAGSGVSEPRNAS 120
 DB 62 lapthlqtelvlrcqketcddlclrvavhlavhgwepedeekfggaadsgvseprnas 121
 QY 121 LQAQVVLFSQAYTARCVLLEVOVPAALVQFGSGVSVVDCFAALGSEVRISYQTQR 180
 DB 122 lqaqvvlfsqaytarcvlllevvpaalvqfgsgvsvvdcfaaigsevrivsyqtqr 181
 QY 181 YKELNHTQQLP-----ALPWLNVSDGDNVHLVNVSEOHFGLSLYWN 225
 DB 182 ykelnhtqqlpcrcglevwnspscwalpwnvsadgdnvhlvlnvseeqhfslslywn 241
 QY 226 QVGGPPKRWKKNLTGP---QITLNTDLVPCICIQVWPLEPDSVR 269
 DB 242 qvggppkprwknlvpppsqv----hshcrpcldckavpygrgslk 284

RESULT 12
 AAEL0920
 ID AAEL0920 standard; Protein; 204 AA.
 XX
 AC AAEL0920;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human gene 12 encoded immune system-related protein HADEX14.
 XX
 KW Human: immune system-related protein; allergy; rheumatoid arthritis;
 KW cancer; Gaucher's disease; viral hepatitis; gene therapy; cytostatic;
 KW diabetes mellitus; arrhythmia; wound healing; ischaemic lesion; AIDS;
 KW Acquired Immune Deficiency Syndrome; viricide; hepatotropic; vasotropic;
 KW autoimmune disorder; inflammation; cardiovascular disorder; hair loss;
 KW wound healing; cell proliferation; skin aging; endocrine disorder;
 KW food preservative.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 50..55
 FT Domain /label= Immunogenic_epitope
 FT 151..168
 FT Domain /label= Immunogenic_epitope
 FT
 XX

PN WO2000166722-A1.
 XX
 PD 13-SEP-2001.
 XX
 PF 07-MAR-2001; 2001WO-US07260.
 XX
 PR 08-MAR-2000; 2000US-187873P.
 PR 11-AUG-2000; 2000US-224367P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ni J, Hilbert D, Kenny JJ, Moore PA, Choi GH, Soppet DR, Ebner R;
 PI Gruber JR, Endress GA, Ruben SM;
 XX
 DR WPI; 2001-589939/66.
 DR N-PSDB; AAD18278.
 XX
 PT Novel isolated immune system-related polypeptide useful for treating
 PT rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease,
 PT diabetes mellitus, arrhythmia, wound healing, ischemic lesions and
 PT viral hepatitis
 XX
 PS Claim 11; Page 312-313; 315pp; English.
 XX
 CC The invention relates to human immune system-related protein and their
 CC DNA. Human immune-system related protein and DNA are useful for
 CC preventing, treating or ameliorating a medical condition in a mammalian
 CC subject, for diagnosing, preventing or treating immune system-associated
 CC disorders, autoimmune disorders (rheumatoid arthritis), inflammatory
 CC disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders
 CC (allergies), infectious diseases (e.g., viral hepatitis), complement
 CC activation disorders, immune complex diseases, neoplastic disorders
 CC (cancer), hyperproliferative disorders (Gaucher's disease), disorders
 CC associated with neovascularisation, diseases at the cellular level,
 CC cardiovascular disorders (arrhythmias), wound healing and epithelial
 CC cell proliferation, endocrine disorders (diabetes mellitus) and
 CC neurological disorders (ischaemic lesions). Immune-system related protein
 CC or DNA is useful for preventing hair loss, skin aging due to sunburn, to
 CC maintain organs before transplantation, to treat weight disorders, to
 CC modulate mammalian characteristics, to change a mammal's mental or
 CC physical state, or as a food additive or preservative. Immune-system
 CC related DNA is useful in gene therapy, for chromosome identification,
 CC radiation hybrid mapping, long range restriction mapping and in forensic
 CC biology. The present sequence represents a human immune-system related
 CC protein of the invention.
 XX
 SQ Sequence 204 AA;

Query Match 29.2%; Score 1079; DB 22; Length 204;
 Best Local Similarity 100.0%; Pred. No. 3.7e-91;
 Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 485 LLYSADDSGFERLVGALASALCOLPLRVAVDLWSRRELSAQGPVAFHAQRRTLQEGGV 544
 DB 1 llysadsgferlvgalasalcolplrvavdlwsrrelsaqgpvafhaqrtrtlqeggv 60
 QY 545 VVLLFSPGAVLCSHNLQGVSGPGAGHDPFRASLSCVLPDFLOGRPGSVYGACFDR 604
 DB 61 vvlflspgavalcsehlqgvsgpgagghdpfraslscvlpdflogrpgsvyvgacfd 120
 QY 605 LLHPDAVPALFRTVPVFTLPSQLPDFLQALQOPRPRSGRLQERAOVSALQPALDSYF 664
 DB 121 llhpdavpalftrtvpvftlpsqlpdfalqalqoprprrsgrlqeraovsalqpaldsyf 180
 QY 665 HPPGTPAPGRGVGPGAGGAGDGT 688
 DB 181 hppgtpapgrgvpgagagdgdt 204
 RESULT 13
 AAAY76048
 ID AAAY76048 standard; Protein; 617 AA.

```

XX AC AAY76048;
XX XX
XX DT 27-MAR-2000 (first entry)
XX DE Murine skin cell protein, SEQ ID NO:303.
XX DE
XX KW Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
XX KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
XX KW secreted; transmembrane; inflammation; cancer; neurological disease;
XX KW angiogenesis; tumour vascularisation; growth disorder;
XX KW developmental disorder; skin wound; hair follicle disorder;
XX KW anti-inflammatory; cytostatic; neuroprotective; vulnery.
XX OS
XX PN Mus sp.
XX PN WO9955865-A1.
XX PD 04-NOV-1999.
XX PF 29-APR-1999; 99WO-NZ00051.
XX PF 29-APR-1998; 98US-0069726.
XX PR 09-NOV-1998; 98US-0188930.
XX PA (GENE-) GENESIS RES & DEV CORP LTD.
XX PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
XX DR WPI: 2000-072177/06.
XX DR N-PSDB: AAZ61753.
XX PT Novel polynucleotides useful for the treatment of various conditions
XX PT including wounds and cancer.
XX PS Claim 4; Page 179-180; 235pp; English.
XX CC The invention relates to novel nucleic acid sequences derived from rat
XX CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
XX CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
XX CC cells. Polypeptides of the invention may be used to treat inflammation,
XX CC cancer and neurological diseases. The proteins may be used to stimulate
XX CC the growth and motility of keratinocytes, to inhibit the growth of
XX CC cancer cells, to modulate angiogenesis and tumour vascularisation, to
XX CC modulate skin inflammation, to modulate epithelial cell growth and to
XX CC inhibit binding of HIV-1 to leukocytes. The invention may also be used
XX CC to treat growth and developmental defects, skin wounds and hair follicle
XX CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded
XX CC by cDNA sequences derived from several mouse, rat or human skin cell
XX CC types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and
XX CC AAY76119 are proteins with an N-terminal signal sequence, indicating
XX CC that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071,
XX CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more
XX CC putative transmembrane domains.
XX SO Sequence 617 AA;

```

```

Query Match 8.3%; Score 305.5; DB 21; Length 617;
Best Local Similarity 25.9%; Pred. No. 4.3e-19;
Matches 130; Conservative 53; Mismatches 179; Indels 139; Gaps 18;

```

```

OY 393 STRAARLG-----EYLLDLOSGGCIQLMDDDLGLM---ACP 427
DB 349 awsdpglqdpmpvyslsqtqgsvpyldlilpflrqencillwrsdvhfawkhvlp 408
OY 428 MDKYIHKRWALVWLACLFEEAALSLILLKKDHAKGWLRLKQDVSGAARGRAALILY 487
DB 409 dd-----apytqlil-----rslsgqtrpylln 434
OY 488 SADDGFERLYGALASALCOLPLRYA-----VDMRSRELSAGPVAWHAORQTL 539
DB 435 aadseagrrilvgalael-----lrlalgggrivdlwgtlvnarigpplwaaterra 489
OY 540 QEGGVVLLFSPGAVALCSEWLQDGVSGPGANGPHDAFPASLSCVLPDGLGRAPGSYVG 599
DB 490 regtlvlllwn-----cagpstacsgdpqaslrlll-----caaprp111 530
OY 600 ACFDRLLHPDAVPALEFRVYVFTLPSQLPDLFGALQ-QPRARPSG-----RLQERAE 650
DB 531 aysfricakgdlprlralpryllrldlprllraldaqpatlasswshlgakrc1knrlr 590
OY 651 QVS-RALQPALDSYFHPGTP 670
DB 591 qchlleaakddygstnsp 611

```

```

RESULT 14
ID AAB55987
ID AAB55987 standard; Protein: 617 AA.
AC AAB55987;
XX 08-MAR-2001 (first entry)
XX DT
XX DE Skin cell protein, SEQ ID NO: 303.
XX KW Mouse; skin cell; cytostatic; anti-inflammatory; anti-HIV;
XX KW neurotropic; neuroprotective; vulnery; immunomodulatory; vaccine;
XX KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
XX KW inflammation; neurological disease.
XX OS
XX PN Mus sp.
XX PN WO200069884-A2.
XX PD 23-NOV-2000.
XX PF 15-MAY-2000; 2000WO-NZ00075.
XX PF 14-MAY-1999; 99US-0312283.
XX PR (GENE-) GENESIS RES & DEV CORP LTD.
XX PA Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD, Murison JG;
XX PI WPI: 2001-007495/01.
XX DR N-PSDB: AAC99686.
XX PT New isolated polynucleotide used in the identification of genetic
XX PT disorders and encoding polypeptides used for treating inflammatory
XX PT disease, cancer and neurological diseases.
XX PS Claim 4; Page 244-245; 352pp; English.
XX CC The present sequence is a polypeptide which is expressed in
XX CC mammalian skin cells. The polypeptide is useful for stimulating
XX CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
XX CC modulating angiogenesis, inhibiting angiogenesis and vascularisation of
XX CC tumours, modulating skin inflammation, stimulating the growth of
XX CC epithelial cells, inhibiting the binding of human immunodeficiency virus
XX CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
XX CC neurological diseases. The polynucleotide can be used as a marker, in
XX CC the identification of genetic disorders, and for the design of

```

FT
/note= "cAMP/GMP-dependent protein kinase

hypersensitive

Experimental demyelinating polyneuropathy

my. Treating a

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: May 19, 2002, 15:19:08 ; Search time 115.29 Seconds
(without alignments)
1032.358 Million cell updates/sec

Title: US-09-608-918-10
Perfect score: 3692
Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPRGVGPAGPGAGDGT 688

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP-archaea.*
- 2: SP-bacteria.*
- 3: SP-fungi.*
- 4: SP-human.*
- 5: SP-invertebrate.*
- 6: SP-mammal.*
- 7: SP-mhc.*
- 8: SP-organelle.*
- 9: SP-phage.*
- 10: SP-plant.*
- 11: SP-rodent.*
- 12: SP-virus.*
- 13: SP-vertebrate.*
- 14: SP-unclassified.*
- 15: SP-virus.*
- 16: SP-bacteriap.*
- 17: SP-archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2554.5	69.2	538	4	Q9BR97	Q9br-97 homo sapien
2	1724.5	46.7	567	11	Q99J43	Q99j43 mus musculus
3	176	4.8	866	4	O43844	O43844 homo sapien
4	172	4.7	866	4	Q96F46	Q96f46 homo sapien
5	166	4.5	864	11	O60943	O60943 mus musculus
6	125.5	3.4	1366	2	Q9Z529	Q9z529 streptomyc
7	120.5	3.3	478	16	Q9RY89	Q9ry89 deinococcus
8	116	3.1	5435	2	Q9L4X2	Q9l4x2 streptomyc
9	115.5	3.1	698	16	Q9PBH5	Q9pbh5 xylella fas
10	115.5	3.1	2424	5	Q9VZ48	Q9vz48 drosophila
11	113	3.1	4928	2	O9ALM3	O9alm3 saccharopol
12	112.5	3.0	1193	16	Q9HZE5	Q9hze5 pseudomonas
13	112	3.0	9510	2	Q93NX9	Q93nx9 streptomyc
14	111.5	3.0	679	2	Q9FP26	Q9fp26 streptomyc
15	111	3.0	361	2	O86440	O86440 pseudomonas
16	111	3.0	620	4	Q9HCN6	Q9hcn6 homo sapien

17	111	3.0	1118	4	Q15426	Q15426 homo sapien
18	111	3.0	1256	11	Q99M76	Q99m76 rattus norv
19	110.5	3.0	593	2	Q9FB15	Q9fb15 streptomyc
20	110.5	3.0	4340	2	O30764	O30764 streptomyc
21	110	3.0	1318	12	P90493	P90493 herpes simp
22	109.5	3.0	2055	4	O75055	O75055 homo sapien
23	108.5	2.9	712	5	Q9BH06	Q9bh06 leishmania
24	108.5	2.9	6797	2	Q9X993	Q9x993 streptomyc
25	108.5	2.9	10917	2	Q93NW6	Q93nw6 streptomyc
26	107.5	2.9	433	16	Q9RTZ2	Q9rtz2 deinococcus
27	107.5	2.9	604	12	Q9YVB7	Q9yvb7 pseudorabie
28	107	2.9	825	4	Q96P01	Q96p01 homo sapien
29	106.5	2.9	274	16	Q9I615	Q9i615 pseudomonas
30	106.5	2.9	430	4	Q96924	Q96924 homo sapien
31	106.5	2.9	884	2	Q9RJ50	Q9rj50 streptomyc
32	106.5	2.9	912	2	Q9K3Y2	Q9k3y2 streptomyc
33	106.5	2.9	1030	4	Q9BY59	Q9bys9 homo sapien
34	106	2.9	1115	4	Q9HD43	Q9hd43 homo sapien
35	106	2.9	4247	2	Q9L8H4	Q9l8h4 streptomyc
36	105.5	2.9	679	16	Q9REE6	Q9rre6 deinococcus
37	105.5	2.9	1060	4	Q9PI27	Q9pl27 homo sapien
38	105	2.8	757	6	Q9N015	Q9n015 macaca fasc
39	105	2.8	766	11	Q9JKQ9	Q9jkq9 mus musculus
40	105	2.8	785	2	Q9K3N5	Q9k3n5 streptomyc
41	105	2.8	921	4	Q9UJ19	Q9uj19 homo sapien
42	105	2.8	3670	2	Q9Z4X5	Q9z4x5 streptomyc
43	104.5	2.8	238	16	O53500	O53500 mycobacteri
44	104.5	2.8	1343	12	Q06635	Q06635 bovine herp
45	104.5	2.8	1505	2	Q9F0D7	Q9f0d7 streptomyc

ALIGNMENTS

RESULT 1

Q9BR97 ID Q9BR97 PRELIMINARY; PRT; 538 AA.

AC Q9BR97; 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 59.1 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOMETRIAL ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC006411; AAH06411.1; -
KW Hypothetical protein.
SQ SEQUENCE 538 AA; 59127 MW; 65E6344DA6A5AFD2 CRC64;

Query Match	69.2%	Score 2554.5;	DB 4;	Length 538;
Best Local Similarity	96.5%	Pred No. 6e-200;		
Matches 475;	Conservative 0;	Mismatches 0;	Indels 17;	Gaps 1;
Qy 1	MPVPWFLLSLALGRSPVLSLERLVGPQDATHCSPGLSCRLWSDILCLPGDIVPAPGPV	60		
Db 1	MPVPWFLLSLALGRSPVLSLERLVGPQDATHCSPGLSCRLWSDILCLPGDIVPAPGPV	60		
Qy 61	LAPTHLOTELVRCKETDCDCLRLRVAVHLAVGHWEPEDEEFKFGGAADSGVEPRNAS	120		
Db 61	LAPTHLOTELVRCKETDCDCLRLRVAVHLAVGHWEPEDEEFKFGGAADSGVEPRNAS	120		
Qy 121	LQAOVWLSFOAYPTARCVELLEVQVPAALVQFGSGVSVVYDCFEAALGSEVRWSYTOPR	180		
Db 121	LQAOVWLSFOAYPTARCVELLEVQVPAALVQFGSGVSVVYDCFEAALGSEVRWSYTOPR	180		
Qy 181	YEKELNHTQOLPALPWLNVNSADGNVHLVNVSEQHFGLSLYWNQVQPPKPRWHKNLT	240		


```
Db 181 YEKLNHTQOLPALPWLNSADGNVHLVNLVSEOHGSLVXNOVGPPKPRWKNLT 240
Qy 241 GQIITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHQNLMQAAARLLTLQSWL 300
Db 241 GQIITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHQNLMQAAARLLTLQSWL 300
Qy 301 LDAPCSLPAEALCWAPGDCQCPVLPPLSWENVTVD-----VNSSE 343
Db 301 LDAPCSLPAEALCWAPGDCQCPVLPPLSWENVTVDKVFEPKLGKHPNLCVQVNSSE 360
Qy 344 KLQLOECLWADSLGPKDDVLLLETRGPQDNRSICALPESGCTSLPSKASTRAARLGEYL 403
Db 361 KLQLOECLWADSLGPKDDVLLLETRGPQDNRSICALPESGCTSLPSKASTRAARLGEYL 420
Qy 404 LDQASGQCQLQWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKG 463
Db 421 LDQASGQCQLQWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKG 480
Qy 464 WLRLKQDVRS 475
Db 481 WLRLKQDVRS 492

RESULT 2
Q99J43 PRELIMINARY; PRT; 567 AA.
AC Q99J43;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE HYPOTHETICAL 62.8 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004759; AA004759.1;
KW Hypothetical protein.
SQ SEQUENCE 567 AA; 62798 MW; CLAAAB79E200681D CRC64;
```

```
Query Match 46.7%; Score 1724.5; DB 11; Length 567;
Best Local Similarity 59.9%; Pred. No. 3e-132;
Matches 348; Conservative 47; Mismatches 125; Indels 61; Gaps 9;

Qy 1 MPVPNFWLSLALGRSPWLRLVGPQDATHCSGLSCLRLWDSILCLPGDIVPAPGPV 60
Db 1 MPVSNFWLLSLALGRNPVVSRLRLMEPQDTARCGLSCLHLWDGDLCLPGLSQAPGPV 60

Qy 61 LAPTHLOTLVLCQKQETCDICLRLVAVHLAVHGHWEPEDEKFGGAADSGVEEPNAS 120
Db 61 LVPTRLQTLVLCQKQETCDICLRLVAVHLAVHGHWEPEDEKFGGAADSGVEEPNAS 116

Qy 121 LQAOVYLSFOAVPTARCVLLEVOVPAALVQFGOSGVVYDCFEAALGSEVRIWSTQPR 180
Db 117 LQAOVYLSFOAVPTARCVLLEVOVPAALVQFGOSGVVYDCFEAALGSEVRIWSTQPR 176

Qy 181 YEKLNHTQOLP-----ALPWLNVSDAGDNVHLVNLVSEOHGSLSLYWN 225
Db 177 YQKELNLTQOLPDCRGLVNRDSTQSCWLPWLNVSTDGDVNLTLVDSQDFSLYLRL 236

Qy 226 QVGGPPKPRWKNLTGQIITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHQNLT 285
Db 237 PVPDALKSLWYKNLTGQIITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHQNLT 296

Qy 286 WQAAARLLTLQSWLWDARCSLPAEALCWAPGDCQCPVLPPLSWENVTVD----- 338
Db 297 WHIARLRLVSPGVWQDAPCCPLGKVTLCWQAPDQSPQPLVPPVPKQKATVNEPQDFQL 356
```

```
Qy 339 -----VNSSEKLOECLWADSLGPKDDVLLLETRGPQDNRSICALPESGCTSL 388
Db 357 VAGHPNLCVQYSTWEKVLQACLWADSLGPFKDDMLLVEMKTGLNNTSYCALEPSGCTPL 416
Qy 389 PSKASTRAARLGEYLQLODQSGQCQLW- DDDLGALWACPMCKYIHKRWALVWLACLLFA 447
Db 417 PSMASTRAARLGEYLQDFRSHOCQMLWDDNMGSWACPMCKYIHKRWALVWLACLLFA 476
Qy 448 AALSILLKLDKDHAKGWLRLKQDVRSAAAGRAALLLYS-----ADDSGFERLVGAL 501
Db 477 AALFEFLLKKD-----RMTSPSPSAACYPIACKGRPAATSG-STSTGCC 522
Qy 502 ASALCQLPLRVAVDLWMSRRELSAOGPVAFWHAQRQTLOEG 542
Db 523 QTTLCT--PPRSASPRSPCRSCR--LSWMHCREAAPLPRG 559

RESULT 3
O43844 PRELIMINARY; PRT; 866 AA.
AC O43844;
DT 01-JUN-1998 (TremBLrel. 06, Created)
DT 01-JUN-1998 (TremBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE IL-17 RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98035683; PubMed=9367539;
RA Yao Z., Spriggs M.K., Derry J.M.J., Strockbine L., Park L.S.,
RA VandenBos T., Zappone J., Painter S.L., Armitage R.J.;
RT "Molecular characterization of the human interleukin (IL)-17
RT receptor."
RL Cytokine 9:794-800(1997).
DR EMBL; U58917; AAB99730.1;
KW Receptor.
SQ SEQUENCE 866 AA; 96122 MW; 88AF626A83F3FF70 CRC64;
```

```
Query Match 4.8%; Score 176; DB 4; Length 866;
Best Local Similarity 21.8%; Pred. No. 9.7e-06;
Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;
```

```
Qy 237 KNLT--GQIITLNTL-----HT---DLVPCICIQWPLEPDS-----VRTN-- 271
Db 66 RNLTSPSSPKDQIQIHLFAHTQOQDLPVVAHIE-WTLQTDASILYLEGAELSVLQNTNR 124

Qy 272 IC---PPREDPRAHQNLMQAAARLLTLQSWLWDAPCS-----LPAAALCWAPGDC 321
Db 125 LCVAFEFSLKLRHHRRWR-----FTFSHFVVDQVEYEVTHLKP-----IPGD 172

Qy 322 PCQP----LVPLPS-----WE-NYTVDNVNSSEKLOECLWADSLGPK 360
Db 173 PNHQSKNLFVPCDCEHARKVVTTPCMSSGSLWDPNITVTEAHQLRVSTLNES----- 227

Qy 361 DDVLLLETRGPQDNRSICALPESGCTSLPSKASTRAARLGEY-----LQDILQ 408
Db 228 THYQILLTSFPH-----MENHSCFEHMHIP-----APRPEFHORSNVTTLRLNK 274

Qy 409 S-----GQCILWDDDL--GALWACP-----MDXYIHKRWALVWLA 442
Db 275 GCCRHQVQIOFFSSCL---NDCLRHSATVSCPEMPDTPPIPDYMPLMWV--WFTIGIS 329

Qy 443 CLLFAAALSLILLKKDHAKGWLRLKQDVR--SGAAA-----RGAALLLSADDS 492
Db 330 ILLVGSVILLIVCMTWRLAGPGESEKYSDDTKYTDGLPLRADLIPPLPKRWIIYSADHP 389

Qy 493 GFERLVGALAS--ALCQLPLRVAVDLWMSRRELSAOGPVAFWHAQRQTLOEGGVVLLFSP 551
Db 390 LYVDVWLKFAQLLTACGTEVALDLEQAISEAGVMTWVGKQKQEMVSNKIIIVCSR 449
```

Db	505	SEVSGDGVDPDLFGAAPRYPLMDRFEEVYFRIQDLEMFQPGRM-HRVGELSG-----DN	55
Qy	663	YFHPPG 668	
Db	558	YLRSPG 563	
RESULT	5		
Q60943			
ID	Q60943	PRELIMINARY;	PRT; 864 AA.
AC	O60943;		
DT	01-NOV-1996	(TREMblrel. 01, Created)	
DT	01-NOV-1996	(TREMblrel. 01, Last sequence update)	
DT	01-DEC-2001	(TREMblrel. 19, Last annotation update)	
DE	INTERLEUKIN 17 RECEPTOR.		
GN	IL17R.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
[1]			
RN	SEQUENCE FROM N.A.		
RP	TISSUE=THYMOMA EL4.		
RC	MEDLINE=96111968; PubMed=8777726;		
RA	Yao Z., Fanslow W.C., Seidn M.F., Rousseau A.M., Painter S.L.,		
RA	Comeau M.R., Cohen J.I., Spriggs M.K.;		
RT	"Herpesvirus Saimiri encodes a new cytokine, IL-17, which binds to a		
RT	novel cytokine receptor.";		
RL	Immunity 3:811-821(1995)		
RL	EMBL; U31993; AAC52357.1; --		
DR	DR		
DR	MGD; MGI:107399; IL17r.		
DR	Receptor.		
SK	SEQUENCE	864 AA; 97807 MW; 343FD51AA687DA31 CRC64;	

Query Match 4.5%; Score 166; DB 11; Length 864;
Best Local Similarity 20.4%; Pred. No. 6.3e-05;
Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;

Qy	192	PALPW----	LNVSADGNVHLVLN----	VSEQHFGLSLYWNOVGPP----	KPRW----	HKNL	239
Db	14	PALGWLLLLLVN	LAPGRASPRLLDPAPVCAQE--	GLSC----	RVKNSTCLDDSWIHPKNL	68	
Qy	240	T-GPOIITLN----	HTDLVPCICIQVWPLEPDS----		VFEN--IC-	273	
Db	69	TPSPFKNIYI	NLVSSTQHGLVPLVHVE-WT	LOTDASILY	LECAELSVLQINFRICV	127	
Qy	274	--PEREDPRAH	ONIWAQARULRLTLQSWL	LDAPCSLPAEA-----	ALCWRAFGPPD----	322	
Db	128	KFOFLSLQ	HHRRKWR-----	FSFSHFVD-----	BQGEYEVTVHHLKRP	IGDPGNHK	176
Qy	323	-----	COPLVPPLS-----	WE-NVTVDVNS	SEKLOLQOECLWADSLG	PLKDDVL	364
Db	177	SKIIFVPC	DESDKMKMTTSCVSSGL	MDPNITVETLOTQHL	RVDFTLWNEST--PYQ----	V	232
Qy	365	LLETRGPQ	DNRSICALPEFSCTSL-----	PSKASTRAAR	LGELYLLODLOGGOCLO	LQWDD	418
Db	233	LLESFSD	SENHS-----	CFDVVKQIFAP	QEEFHQRANVTFLSK	KFH-----	274
Qy	419	DLGALWAC	-----		PMDKYTHKSWA----	LV	439
Db	275	-----	WCCHHHVQVOPFSSCL	NDCLRHAVTVPCPVIS	NTVPKVADYI--PLWVTGLY	328	
Qy	440	WLACLLFAA	LSLILLKHKHGWLLK	LKODVRSGA-----		476	
Db	329	LIAILLVGS	VILIICT-----	W-RL-----	SGADQEKHGDSKINGIL	PVADLTTP	375
Qy	477	AARGRAALL	YSADDSGFERLVGN	LASAL--COLPLRV	AVDLWSRRELSAOG	PVAWEHAQR	535
Db	376	PLRPRK	YVIVYSADHPLYVE	WLKFAQFLITAC	GTEVALDLLEEQVIE	SGVMTWVSROK	435
Qy	536	ROTLQEG	GVVLLFSPG-----		AVAL--CSEWLOD	GVSGPAHGPHDAFRAS	580

```
Db 436 QEWESNSKIIILCSRGTOAKWAILGWAEPVQLRCDHWKPG-----DLFTAA 485
QY 581 LSCVLPDFQGRAPSGYSGACFDRLHDPALFTRVPTLPSQLPDFLQALQPPRAP 640
Db 486 MNMILPDFKRPACFGYVVCYFSGICSDRDPDLFNITSRYPLMDREVEEYFIQDLEM 545
QY 641 RSGRLQERAEQVSRLQALPDLSYFHP 667
Db 546 EFGRMHVRELATG-----DNYLQSP 565

RESULT 6
Q92529 ID Q92529 PRELIMINARY; PRT: 1366 AA.
AC Q92529;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE PUTATIVE LARGE PRO/ALA/GLY-RICH PROTEIN.
GN SC9F2.06C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrett B.G., Rajandream M.A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL035559; CAB37473.1; -
SQ SEQUENCE 1366 AA; 141042 MW; FB954569D87D029B CRC64;
```

Query Match 3.4%; Score 125.5; DB 2; Length 1366;
Best Local Similarity 23.2%; Pred. No. 0.23;
Matches 177; Conservative 58; Mismatches 248; Indels 281; Gaps 43;

```
QY 21 LERLVGPQDATHCPSGLSRLWSDILCLPGDITVAPGPVLAATHQTELVLRCQKETDC 80
Db 675 LDGALADPDAGH---PLTIRLLSEVRAALFGP--PAPVPVTRDAVFTAYLDL----- 721
QY 81 DLCLRVAVHLYVGHVEEPEDEKFGAA-----DSGVEPRNASL-QAOWVLSFOA 131
Db 722 -MCLRVATRIA-----DENLGRGTVRRLAAKVSQGVHEAARRSLGPGQGLDRES 771
QY 132 YPT-----ARCVLLE-VQVPAAL-----VOFGQSVG 156
Db 772 FETLPCGPAPARLGGTGWAPVLAEGLVPTGSGYRFAHEELADWIOQTHIDLGEALR 831
QY 157 SVVYDCFEAALGSEVRINSYTOPRYEKELNHTQOLPALPWLNVYSDAGDNVHLVNSEEQ 216
Db 832 ALVHR-RDTPLGT-----HTRTLP-VPHRI---GSVVEALLLLA-RQ 868
QY 217 HGLSLYNQVQGPKPVRHKNLTGQITLNTDLVPLCICIQVWPLEPDSVRTNICPFR 276
Db 869 H-----GVPO-----LALTLEELVHAL-----D 886
QY 277 EDPRAHQLNWQAARLRLTLQSWLLDAPCSLPAEAL-----CWRAPGDPQCPVLP 329
```

```
Db 887 RDPHS-----WAAWL-----LAELTRVPDAPYTDVLLADGIAERAGDQPTQVFGP 938
QY 330 LSWENVTVDVNSSEKLQIQECLWADSLGPKDDVLLLETRGPQDNRSLEALEPSGCTSLP 389
Db 939 AFWTAPRVPATRLDL-LRRLVLAD-----GPPH-----EPG-----P 970
QY 390 SKASTRAARLGEYLLQDLSQGCQLQ-LWDDDLGALWACPMCKYIHKRWALVWLACLLFAA 448
Db 971 RHLDTAAG-----LLVADPRTVQPLLVRFWDFDERPLPATP-----HATVA-----TA 1012
QY 449 ALSILLILLKDHAKGWLRLIKQDVRSAAAGRAALLIYSADDSFERLVGALA-----SA 504
Db 1013 AQALLHTRHRLDGLTEVLVDSTHRA-----DELLAVLABEEPSA 1054
QY 505 LCQLPLRVAVDLWSRRLSQAQPGVAWFHAQR-----OTLQEGGVVLLFSP----- 551
Db 1055 LCR-----AVERWARDERPAHRAAVTHGLTAPHARGADFTLLRHAALVLLAGPSDP 1109
QY 552 ---GAVALCSEWLQDVGSGGAGHP-----DAFRA-----SLSCVLPDFLQGRAP 595
Db 1110 LRGGALALL-----VQDPDRHRLPALDLFAACDPLPPSAVAALPTH-----PE 1157
QY 596 SYVCACFDRLHDPALFTR-----TVPVFTLPQLPDFLQALQPPRPSRQLQERAEQ 651
Db 1158 PVLEAFRALRILGPDAGEALRRLADATTALT--HRVALVGRVTVTERPETAGHL---AAY 1212
QY 652 VSRL--OPALDSYFHPPT-----PAPRGVGVGPGAGDG 687
Db 1213 VDRRLDRPAPRAVLLPLVTRLLDDGPEPARAA--LAGVLAADG 1254

RESULT 7
Q9RY89 ID Q9RY89 PRELIMINARY; PRT: 478 AA.
AC Q9RY89;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE HYPOTHETICAL 51.2 KDA PROTEIN.
GN DR0061.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AF001869; AAF09658.1; -
DR TIGR; DR0061; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 478 AA; 51231 MW; 5E74CDEB8FBA5C95 CRC64;
```

Query Match 3.3%; Score 120.5; DB 16; Length 478;
Best Local Similarity 24.3%; Pred. No. 0.15;
Matches 99; Conservative 39; Mismatches 148; Indels 121; Gaps 17;

```
QY 276 REDPRAHQLNWQAARLRLTLQSWLL-----DAPCSLPAEALCWRAPGDPQCPV 327
Db 110 RADPRERPLLEAAQ-----RIRRWLARLARAPGLPTVPAHSVLY-----AQEASALS 161
QY 328 PPLSWENVTVDVNSSEKLQIQECL--WADSLGPKDDVLLLETRGPQDNRSLEALEPSG 384
```

```
Db 162 APAAGDGLTHLGDPDRLAHQGLRLSWR-SLGPWNQWLLVQDETSPQVSGGQALLRP-- 218
QY 385 CTSPLSKASTRAARLGEYLLQDLSQGCQLWDDDLGALWACPMCKYIHKRWALVWLACL 444
Db 219 --DLPASERQLFTVGGQQLVLFSG----- 242
QY 445 LFAAALSILILLKKDKHAKGWLRLKQDVRSGAAARGRA-ALLLYSADDSGFFERLVGALAS 503
Db 243 -----DYVLLRRRADAQTHLAR-----LAALGRACALLLLPAEQHGRRLRLARTLAR 290
QY 504 ALCOLPLRVAVDLWSRRELSAOGPVAVWFHQAORROTLOEGGVVLLFSPGAV----- 554
Db 291 RLRGDPPR-ADDPASGQVAFATACLAARR-SLQRLEPLLLGRFSPAQAHAHEAQAQA 348
QY 555 -----ALCSEWLQDVSGP--GAHGPHDAFRASLSCVL-----PDFL 589
Db 349 LLHLPPHYAAQVQAAEHAASAEPLREAAATDPLVGREANGPVLPATGHFIYFPLGPDPL 408
QY 590 QGRAPGSVVGACFRL--LHPDVPALPRTVPVFTPLPSQLPDFLAL 634
Db 409 EVRLPG-----DRLLTLRPD-----YRAELVAVLPQAAVTVGDL 443

RESULT 8
Q9L4X2 ID Q9L4X2 PRELIMINARY; PRT; 5435 AA.
AC Q9L4X2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NYSJ.
GN NYSJ.
OS Streptomyces noursei.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1971;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11455;
RX MEDLINE=20334850; PubMed=10873841;
RA Brautaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,
RA Valla S., Zolchev S.B.;
RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
RT Streptomyces noursei ATCC 11455: analysis of the gene cluster and
RT deduction of the biosynthetic pathway.";
RL Chem. Biol. 7:395-403(2000).
DR EMBL; AF263912; AA071767.1;
DR HSP; P25715; IMLA.
DR InterPro; IPR001227; Acyltransf_domain.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Phosphopant_attach.
DR Pfam; PF00698; Adh_zinc; 3.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF02801; ketoacyl-synt; 3.
DR Pfam; PF02801; ketoacyl-synt; 3.
DR PROSITE; PS00075; ACP_DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 5435 AA; 562659 MW; AA55465DF087A38C CRC64;
```

Query Match 3.1%; Score 116; DB 2; Length 5435;

Best Local Similarity 24.8%; Pred. No. 8.3;

Matches 118; Conservative 44; Mismatches 196; Indels 118; Gaps 27;

QY 290 RLRLTLTQSWLLDAPCSLPAEAALCWAPGDPCCOPLVPPLSWENVTVDSSEKILQOE 349

Db 2817 RVRRLAV-SHAFHSPMLDPAFRAVAEGLEYHQPRIPVVS--NVTGEVAAAEELCAAD 2873

```
QY 350 CLWADSLGP---LKDDVLLLETRGPQDNRSICALPESGCTS-----LPSKASTRAARLG 400
Db 2874 -YVVRHVRAFVRFDAGVVRTLAERGA---TAFLEIGDPCVLSALARGVLPALVPTPT--- 2926
QY 401 EYLLQDLSQGCQLWDDDLGALWACPMCKYIHKRWALV-WLACLLLFAAALSLL-LKK 458
Db 2927 --LRKORDEESAL-----LAGL-----ARLHVAGVTVVDSAAALTTGTGARGTDLPTIYAF 2972
QY 459 DHAKGWLRLKQDVRSGAAA-----RGRAALLLYSADDSGFFERLVGALASALCOL 508
Db 2973 QREYWPPELAEPAGGADAADAEFWAAVERADATATAAHLIDIGDQ--LGAVLPALSAW 3030
QY 509 PLRVAVDLWSRRELSAOGPV---AWFHAORROTLOEGGVVLLFSPCAVALCSEWLQDG 564
Db 3031 RTR-----RRTTSATNALRHRESWEPLSLAGTPHTGGVLVLV---PAAATDPMWADV 3080
QY 565 VS--GPCAHG---PHDAF-RASLSCVLPDFLOGRAPSYYG-ACFDRLLHPDVP----- 612
Db 3081 VAALGPDARRVDPADGTDRAALAAALTEAADTAPAVVSLIALDETSGDDAVPAGTTA 3140
QY 613 -----ALFRT---VPVFTLP-----SOLPDFLGAQOPRAPRS-----GRL----- 645
Db 3141 TAALVQALADTGAPAPLWALTRGVAALPD-----EQTAPAAQAAVWGLGRIALALEPRH 3195
QY 646 -----QERAEQSVSRALQPAL-----DSYFHPGPTAPGPGVGPAGPGAGDGT 688
Db 3196 WGLGLVDLPADLDETRRLPAALADAGDEQDLALRATGAYGRRITPAPADDPAGT 3251

RESULT 9
Q9PBH5 ID Q9PBH5 PRELIMINARY; PRT; 698 AA.
AC Q9PBH5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN XF2169.
GN XF2169.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carver H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorfy H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marinho C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.T.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
```

RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL: AE004030; AAF84968.1; .
DR InterPro: IPR001440; TPR.
DR Pfam: PF00515; TPR; 6.
DR SMART: SM00028; TPR; 3.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 698 AA; 77588 MW; 91326560086F5A7F CRC64;

Query Match 3.1%; Score 115.5; DB 16; Length 698;
Best Local Similarity 21.8%; Pred. No. 0.63;
Matches 160; Conservative 90; Mismatches 234; Indels 251; Gaps 43;
Qy 61 LAP-----THLQTVLRCCKE--TDCDLCLRVAVHLAVHGWPEDEKPG-----G 107
Db 69 LAPERPELHLGLLGLLAQDLGSAALSRTRL-----DPNPFNTYVMOAHLA 118
Qy 108 AADSGVEPRNASLQAVVLSFQAVP-----TARCVLLEVVQVPAALVQGVQSV-----G 156
Db 119 TARGALDAEQLSRATRIIS--DHPQLLVNATIALRRGQVDOALTMLSQAVELLPEEP 176
Qy 157 SVVYDCFEAALGSVRINWYTPRY-----EK-----ELN-HTQQLPALPWLNVASADGN 205
Db 177 AVLE-----SUG-----FAYLQKGHATAFAERAFORVIELNPHVIPRYALIAQLAQRG-R 225
Qy 206 VHLVNLVSEEHQFGLSYLVNQVGGPK-----PRWKNLTGPQITLNTDLVPCLCIQV-- 260
Db 226 LDDALRILEE-----ALSL-----PEGDTPALHR-LTGEFELLAGHPDRALTHLRQVLA 273
Qy 261 -WPLEPDSVRTNICFRE-----DPRAHNLWQAARLR----- 293
Db 274 TWPEPDRLEALLTAWKOLDMDDDARITLDAALDIKPRNH-DLW-LARLAVAPVGSDEAR 331
Qy 294 LTLQSWLLDAPCSLPAEAALCWRAFGDPCQPLVPPLSVENVTVDVNSSEKLQLECLWA 353
Db 332 IVIEWLSAMPEHLPALETLM----- 352
Qy 354 DSLGLPKDDVLLLETRGPQDNRSICALEPSGCTSLFSKASTRAARLGEVILQ--DLQSQG 411
Db 353 -SLHDIQNNPEAAETVA---RQIVAEPRGLSG-----EQRIVEALLQRPPTAVA 399
Qy 412 CLQLWDDDLGALWACPMKDIHKRWAL-VWLACL-----LFRAAALSLILLKKDHAKGWL 465
Db 400 CLQQLIESVPE-----HERITLREWGLVQDRAGQFEAALATWLQFAEQAKYRL 449
Qy 466 RLLKQ-----DVRGGAARGAALLLYSADDSGFERLVGALA---SALC----- 506
Db 450 PLPQTWTNSKQWPDLAIPAELARPLLIWPGPGSHVETVMMWASMLCADRYSET 509
Qy 507 -----QLPLRVAVDLMSRRELSAQGP-----VAWFH-----AQRR 536
Db 510 PPADPLQRYETVSELTSGTLTPQALIDAW-RQQLPARGIEDGNVIDWLWWDNSLTLALR 568
Qy 537 QTLQEGGVVLLFSGAVALCSEWLQDGVSGCANGPHDAFRASLSCVLPDFLQGRAPGS 596
Db 569 PHLPEGRLLIIVLRDPRMLL--DWIAYGSPILP---LDSLQQAANW-LGDIIL-----N 616
Qy 597 YVGACFDRLHP-----DAYPALFRTV-PVFTLPSQLPDFLQALQOQPPRPSGRL 645
Db 617 QIAALHEDLPHHILRLDGDIEDNPQALATTLEDIFGSPFPIPP-----SLEAPRLP-AGRW 672
Qy 646 QERAEQVSRALQPAL 660
Db 673 RDYREVLSAFAVIL 687

RESULT 10

Q9VZ48

ID Q9VZ48

AC Q9VZ48

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

Query Match

Best Local Similarity

Matches 128; Conservative

Qy 269 RTNICPFREDPRAHQLWQAALRLR-----LTLQSWLLDAPCSLPAEAALC----- 314

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CG2174 PROTEIN.
GN CG2174.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.C., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,

RA Brandon R.G., Rogers J.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Avril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,

RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saudefs R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RL "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

EMBL: AE003484; AAF47980.1; .

HSSP: P08799; 1MND.

FlyBase: FBgn0030252; CG2174.

InterPro: IPR000048; IQ.

InterPro: IPR001609; myosin_head.

InterPro: IPR000857; MyTH4.

Pfam: PF00612; IQ; 3.

Pfam: PF00063; myosin_head; 1.

Pfam: PF00784; MyTH4; 1.

PRINTS: PR00193; MYOSINHEAVY.

PRODOM: PD000355; myosin_head; 1.

SMART: SM00015; IQ; 3.

SMART: SM00242; MYSC; 1.

SMART: SM00139; MyTH4; 1.

SQ SEQUENCE 2424 AA; 267616 MW; 8AD62AA33F9AA5D0 CRC64;

Query Match 3.1%; Score 115.5; DB 5; Length 2424;

Best Local Similarity 25.3%; Pred. No. 3.2;

Matches 128; Conservative 33; Mismatches 169; Indels 175; Gaps 27;

Db 906 RRRVQOLRSDHRRQREAREAREAKQAVLERSQLSYLD----IPAELAFIYSKL 961
 QY 315 --WRAPGPGCPQPLVPPLSWENTVDVNSSEKIQLOECLWADSLGPKD---DVLLETR 369
 Db 962 QGWSPPHGD--RHLVRVLG-----TVPGPPSAVOLPEDLGQFSFGKSSVYCNGLRQPR 1015
 QY 370 GQDNRSICALPSGCTSLPSKASTRAARLGEVLLQDLSQCCLQDWDLLD---GALWAC 426
 Db 1016 -----REP-----ITAPLLTRAASRDQDFQDALAVFKLRLWSNDKALEGAKEKL 1060
 QY 427 PMDKYIHKR-----WALV--WLACLLFAAALSILI 453
 Db 1061 LADYIVHKALSSRGLRDEILVOLCNQVHLPNNSGEATRLWQLLQCLCCFCQPSAASFY 1120
 QY 454 LL-LKKDHAKGWLJR--LLKODVR-----SGAAARGRA-----ALLYIADDSGFERLV 498
 Db 1121 LMRFVDEAPESLRPLLLRQLLQOQGGTSSGAVGAGACRSFVPAWLEWRAWTRGCD--- 1177
 QY 499 GALASALCOLPL-----RVAVDLWSRRELSAQGPVAFHQAQRRTLOEGGVVLLFS 550
 Db 1178 -----MALPLTLPDEASQTVAVDSWTSCEAAALAVSSLGVASR-----GWTLVL--- 1222
 QY 551 PGAVALCSEWLQGVSGGAGHPDAFRASLSCVL-----PDFLQGRAPGSY 597
 Db 1223 -----DQOQLTDSGLDYVMDLIAEKLCLCPAFPAPRSDLLR 1259
 QY 598 VGACFERLLHPDAV--PALFTVTVFTLPSQLPDFLQALQOQPRAPRSRGLQERAEQVSR 655
 Db 1260 SGAKFARTLPDAVKRPA---VP-----PPAPPTSSGKEDVPRERRSSR-----ELLSRS 1306
 QY 656 LQPAL-DSYFHPHPPGTPAPRGVCPG 679
 Db 1307 --SALNERYFEREPSPGPGQSGSTG 1329

RESULT 11

QYALM3 ID Q9ALM3 PRELIMINARY; PRT; 4928 AA.
 AC Q9ALM3;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE POLYKETIDE SYNTHASE EXTENDER MODULES 5-7.
 GN SPND.
 OS Saccharopolyspora spinosa.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Pseudonocardineae; Pseudonocardiaceae;
 OC Saccharopolyspora.
 OX NCBI_TaxID=60894;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21257765; PubMed=11358695;
 RA Waldron C., Matsushima P., Rosteck P. R. Jr., Broughton M.C.,
 RA Turner J., Madduri K., Crawford K.P., Merlo D.J., Baltz R.H.;
 RT "Cloning and analysis of the spinosad biosynthetic cluster of
 Saccharopolyspora spinosa."
 RL Chem. Biol. 8:487-499 (2001).
 DR EMBL; AY007564; AAG23263.1; .
 DR HSSP; P25715; 1MLA
 DR InterPro; IPR001227; Acyltransf_domain.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Phosphopant_attach.
 DR Pfam; PF00698; Acyl.transf. 3.
 DR Pfam; PF00109; ketoacyl-synt. 3.
 DR Pfam; PF02801; ketoacyl-synt_C; 3.
 DR PROSITE; PS50075; ACP_DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
 DR PROSITE; PS0012; PHOSPHOPANTETHEINE; 3.
 KW Phosphopantetheine; Transferase.
 SQ SEQUENCE 4928 AA; 517247 MW; C22C6361888F865A CRC64;

Query Match

3.1%; Score 113; DB 2; Length 4928;

Best Local Similarity 21.7%; Pred. No. 13;
 Matches 147; Conservative 66; Mismatches 238; Indels 226; Gaps 31;
 QY 104 RFG--GAADSGVEEPNRLAQVVLSTARCVLLEVOVP-AALVQFGQSGSVY 160
 Db 4128 EFGPDGALSALVEQCLAGSDQAGRV--AAIPLMRRDRDEVTAVAAALAHVHRGAVDM 4184
 QY 161 DCFEALAGS---EVRIWSTQPRYEKELNHTOOLPALPWLNVSDAGDNVHLVNLVSEEQH 217
 Db 4185 SACFAGTGARTVELPTAFQORY-----WLAGQADGCGDGVVADPVDAR- 4229
 QY 218 FGLSLYWNQVQ--GPKPRWHKNIQTPIITLNTDLVPCLCIQVWPLEPDSVYRNICPDR 276
 Db 4230 -----FWELVERADPEP-----LVDELCI-----DRDQ-----PFR 4255
 QY 277 E-----DPRAHONLWO-----AARLRLLTLOSLLDADPCSLPAEAA 312
 Db 4256 EVLPVLASMRKQROEALADSWRYQVRWSVEVPSAALRGV---WLVLVLPADVPRD-- 4309
 QY 313 LWRAPGGDPCQP--LVPLSWENTVDVNSSEKIQLOECLWADSLGPKDVLLETRG 370
 Db 4310 -----QPAVVIDALIARGAEVAVLELQDQORSALVDKVRAVIAD--RTEVTG 4356
 QY 371 PQDNRSICALPSGCTSLPSKASTRAARLGEVLLQDL-QSGCLOLWDDDLGALWACPM 429
 Db 4357 V---LSLLAMDGMPCAAHPHLSRGVAIV--ILTQVIGDAGVSAPLWLATTGGVEAGTD 4411
 QY 430 KYIHKRWALWACLILFAAALSILLLKKDHAKGWLRLKQDVRSGAAARGRAALLLYSA 489
 Db 4412 GPADPDHGLIW-----GLGRVVGLEHPQWNGGLID-----LPETL 4446
 QY 490 DSGFERLVGALASALCOLPLRV-AVDLWSRREL-SAQGPVAFHQAQRRTLOEGGVVYL 547
 Db 4447 DETSRNGLVAALAGTAAEDQLAVRSSGLFVRVRAARNP-----RSETWRSRGTVLI 4499
 QY 548 LFPSCAV-ALCSEWL-----QDGVSGPGAGH-----PH 574
 Db 4500 TGGTGALGAEAARWILARRGAELHLVLSRRGPEAPGAADLGAELTELGVKVTVLACDVDR 4559
 QY 575 DAFTASLSCV-----LPDFLQGRAPGSYVGACFDRL 606
 AC Q9HZE5; PRELIMINARY; PRT; 1193 AA.
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE HYPOTHETICAL PROTEIN PA3063.
 GN PA3063.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,

DE AMPHI.
GN AMPHI.
OS *Streptomyces nodosus*.

RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases


```
RN RP SEQUENCE FROM N.A.
RC STRAIN=97000351; PubMed-8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RX Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb streptomycetes coelicolor A3(2) chromosome.";
DR EMBL; AL442120; CAC09545.1; -
DR InterPro; IPR002296; N12N6_mtfase.
DR PRINTS; PR00507; N12N6MTFRASE.
KW Hypothetical protein.
SQ SEQUENCE 679 AA; 72161 MW; 514562165037F451 CRC64;

Query Match 3.0%; Score 111.5; DB 2; Length 679;
Best Local Similarity 23.8%; Pred. No. 1.3;
Matches 120; Conservative 43; Mismatches 177; Indels 165; Gaps 30;

Qy 294 LTQLSWLDAPCSLPAEALCWAPGGDCQPLVP-----PLSWENVTVDV 339
Db 56 LRAQKLAIEV---PRERV--WQLAGHPGPVAAVLQAGCVLLIHDRPTLW--LDASA 108
Qy 340 NSSEKLQLOECLWADSLGPKDDVLLLETR---GPQDNRSCLALEPSGCTSLFSKASTRA 396
Db 109 GSDERL-----ADLLPALDOV--LDARFGTGPQ-----RAVTTAGPRLPLSAPLLRG 155
Qy 397 AR-----LGEYLQDLSGQCQLQDWD--DLGALWACPMCKYTHKRWALV 439
Db 156 LAELAAGPGARKAYEFLLGRHL--DANPROYTLTPDPLADLMAELAGPARTVLDP----- 208
Qy 440 WLAC-----LLFAAALS-----ILLKKDHAKGWLRLKODVRSAGAAAGRAALLLYSAD 490
Db 209 ---ACGTGSLRAAAATTPGQELYGOESDPAALATLRLALSTDAVR-----IAAGD 260
Qy 491 DSGFERLVGALA-SALCQLPLRVAVDLWSRRELS-----AQGPVAVF-HAQR 536
Db 261 SLRADARTGLRADAALCHPPFNER--NWGHDELAYDPWEYGFPPARTSELAWVQHALAR 318
Qy 537 QTLQGGVVVLLFSCAVA-----LCSEWLQDG-----VSGP-GAHPGHDAFRASLSCV 584
Db 319 ---VRDGGTVVLLMPAAASRRSGRRVRLDLRRGALHVAIALPVGAAPPYN----- 367
Qy 585 LPDFL-----QGRAP-----GSYVG-----ACFDRLLHPDA 610
Db 368 LPLHLWLRPERAPAQPCVLLADTGQFAGEGGPDWRSVRDAVLDATWAFDRTGRLDD 427
Qy 611 VPALFRTVPVFTLPSPDLGLAQOQPRAPSG---RLQERAEQVSRALQPALDSYFHPP 667
Db 428 RPGLARCLPVIELDDDDVDLAPARHLPPPTATGGTERLTDVRELRDTELRLTAEL---AP 484
Qy 668 GTPAPGRG-----VGPAGPGA 684
Db 485 AAPGTGRAALRRPLTTVGELARGGA 509

RESULT 15
O86440 PRELIMINARY; PRT; 361 AA.
AC O86440;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE XCPY PROTEIN.
GN XCPY.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN RP SEQUENCE FROM N.A.
```

```
RC STRAIN=WCS358;
RA de Groot A., Gerritse G., Lazdunski A., Filloux A.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; X81085; CAA56986.1; -
SQ SEQUENCE 361 AA; 40565 MW; C364A65E526F998E CRC64;

Query Match 3.0%; Score 111; DB 2; Length 361;
Best Local Similarity 21.4%; Pred. No. 0.62; Indels 194; Gaps 24;
Matches 92; Conservative 31; Mismatches 113;

Qy 192 PALPWL-----NVSADGONVHLVNVSEEQHFLSLY 223
Db 10 PARFWLLLRPGSPDWLLVEGGIAQROGQGEPPANLQA---RVALIVAEHCSHF----- 61
Qy 224 WNVQGGP---KPRWHKNLTGPQII---TLNHTDLVPCLCIQVWPLEPDSVRTNICPFRE 277
Db 62 --QLPAPPGLRREW-----PLLEDRLQSADEVVCGCLG-----RE 97
Qy 278 DPAHQNLWQAARLRLTLQ-----SWLLDAPCS-----LPAE---AALCWR 316
Db 98 -----AGQVRLLTVAROHLGWRGQCCKVKGVSVECCWAEFQLLPAPGPGVAVCMQ 147
Qy 317 APGDPQCPQLVPP-----LSWENVTVDVNSSEKLOLQ---ECLWADSLGPKDDVLL 366
Db 148 KPGFSLMKGLSEGAGERWLAWPDLGVVRPEPQLRVPLEGRWPDELAPLDSLPGLF 207
Qy 367 ETRGPQDNRSICALEPSGCTSLPSKASTRAARLGEYLLQLDLSGQCLOL---WDDDLGAL 423
Db 208 EORRP---RTFVG-----PSRAQRQ-----LL---AACLLLLACTW---GGL 239
Qy 424 WACPMCKYIHKRWALVWLACLLFRAALSILLKKDHAKGWLRLKQDVRSGAAARGRAA 483
Db 240 W-----LAQWRQVQL-----W-----QAQ 254
Qy 484 LLLYSADDSGFERLVGAL-----ASALCQLPLRVAVDLWSRELSAOGPVAV-----FH 532
Db 255 VLAVTGEQAGPRQAAQALKRLREAELOQQLRVRLADLQGRLOAWLQAHPGWRLOAVRFD 314
Qy 533 AQRQTLOEQ 542
Db 315 GQRWHVRVEG 324

Search completed: May 19, 2002, 15:19:16
Job time: 6752 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 14:10:14 ; Search time 43.31 Seconds
(without alignments)
388.012 Million cell updates/sec

Title: US-09-608-918-10
Perfect score: 3692
Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPGRGVGPGAGDGT 688

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	305.5	8.3	617	4	US-09-188-930-303
2	176	4.8	866	2	US-08-620-694A-10
3	176	4.8	866	3	US-09-022-255-10
4	176	4.8	866	3	US-09-022-696-10
5	176	4.8	866	3	US-08-978-773-4
6	176	4.8	866	3	US-09-022-253-10
7	176	4.8	866	3	US-09-022-260-10
8	176	4.8	866	4	US-09-022-259-10
9	176	4.8	866	4	US-09-022-257-10
10	166	4.5	864	2	US-08-620-694A-2
11	166	4.5	864	3	US-09-022-255-2
12	166	4.5	864	3	US-09-022-696-2
13	166	4.5	864	3	US-08-978-773-2
14	166	4.5	864	3	US-09-022-253-2
15	166	4.5	864	3	US-09-022-260-2
16	166	4.5	864	4	US-09-022-259-2
17	166	4.5	864	4	US-09-022-257-2
18	147	4.0	330	4	US-09-188-930-125
19	126.5	3.4	1711	2	US-08-342-930-2
20	113	3.1	4928	4	US-09-036-987A-5
21	113	3.1	4928	4	US-09-370-700-5
22	109.5	3.0	428	4	US-09-423-340-4
23	104	2.8	3724	2	US-08-804-227C-10
24	104	2.8	3724	2	US-08-804-198-4
25	103	2.8	6095	4	US-09-144-085-2
26	102.5	2.8	1785	4	US-09-341-587-3
27	102	2.8	1208	4	US-09-463-702A-2

```

28 102 2.8 2629 2 US-08-751-189-4 Sequence 4, Appli
29 102 2.8 2629 2 US-09-060-836-4 Sequence 4, Appli
30 102 2.8 2629 4 US-09-184-445-4 Sequence 4, Appli
31 101.5 2.7 469 3 US-08-753-007A-8 Sequence 8, Appli
32 101.5 2.7 469 4 US-09-398-496-8 Sequence 8, Appli
33 100.5 2.7 405 2 US-08-483-151-4 Sequence 4, Appli
34 100 2.7 5588 4 US-09-036-987A-6 Sequence 6, Appli
35 100 2.7 5588 4 US-09-370-700-6 Sequence 6, Appli
36 98.5 2.7 647 3 US-08-753-007A-32 Sequence 32, Appli
37 98.5 2.7 647 4 US-09-398-496-32 Sequence 32, Appli
38 98.5 2.7 890 1 US-08-445-640-2 Sequence 2, Appli
39 98.5 2.7 890 3 US-08-170-558-2 Sequence 2, Appli
40 98.5 2.7 890 3 US-08-447-314-2 Sequence 2, Appli
41 98.5 2.7 890 3 US-08-445-461-2 Sequence 2, Appli
42 98.5 2.7 911 1 US-08-286-305A-1 Sequence 1, Appli
43 98.5 2.7 911 2 US-08-441-104A-1 Sequence 1, Appli
44 98.5 2.7 911 2 US-08-440-816A-1 Sequence 1, Appli
45 98.5 2.7 911 4 US-09-417-381A-1 Sequence 1, Appli

```

ALIGNMENTS

```

RESULT 1
US-09-188-930-303
; Sequence 303, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 303
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-303

```

```

Query Match 8.3%; Score 305.5; DB 4; Length 617;
Best Local Similarity 25.9%; Pred. No. 2.6e-20;
Matches 130; Conservative 53; Mismatches 179; Indels 139; Gaps 18;

```

```

QY 237 KNLGPGQIITLNTHTDLVPCLCIQVWPLEPDSVYRTNICPFREDPRAH-QNLWQAAR-----290
DB 183 KIVSGGHVVDLYEFLLPCMCIEASVYLDQETVRRKKCFQSWPEAFYGSDFWQSIRTDYS 242
QY 291 -----LRLLTQSLWLDAPCSLPAEALCWAPGDCQCPVLPPLSWEN-----VTVDV 339
DB 243 QHQNVMALTLR-----CPLKLEASLCWRQDPLTPCETLPNATAQEGSEGWILENVDL 295
QY 340 N-----SSEKILQEQELWADSLGPKDDVILLLETRGPDNKRSLCALPEPGCTSLPSKA 392
DB 296 HPOLCFKFSFENSSHVEC--PHQSGSLPSWTYSMDTQAQ-----LTLHFSRRTYATFSA 348
QY 393 STRAARLG-----EYLLQDLQSGQCLOLWDDDLGALW---ACP 427
DB 349 AWSDPGLGPTDPMPPVYSISQTQGSVPVTLDLIIPLFRQENCILVWRSDVHFAWKHVLCP 408
QY 428 MDKYIHKRWALVWLACLLFAAALSLLILKKDHAKGWLKLLKQDVPSGAAARGRAALLY 487
DB 409 DD-----APYPTQLL-----RSLGSGRTRPVLLJH 434
QY 488 SADDGGERLYGALASALCQLPLRVA-----VDLWSRRELSSAOGPVAMFHAORROTL 539

```

Db 435 AADSEARLLVGAEL-----LRTALGGGRDVIIVDLWEGTHVARIGLPLWLAARVA 489
QY 540 QEGGVVVVLFSPGAVALCSEWLQDGVGPGAHGPHDAFRASLSCVLPDFLOGRAPGSYVG 599
Db 490 REQQTLLWN-----CAGPSTACSGDPOAASLRTLL-----CAAPRPLLL 530
QY 600 ACFDRLLHPDAVPALEFRTVPVFTLPSQLPDLFLGALQ--QPRAPRSG-----RLQERAE 650
Db 531 AYSRLCAKGDIDPRPRALPRYLLRLDLPRLLRALDAQPATLASSWSHLGAKRCLKNRLE 590
QY 651 QVS-RALQPALDSYFHPGTP 670
Db 591 QCHLLEAKDDYQGSTNSP 611
RESULT 2
US-08-620-694A-10
; Sequence 10, Application US/08620694A
; Patent No. 5869286
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694A
; FILING DATE: 21 MARCH 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; CLASSIFICATION: 435
; FILING DATE: 23 MARCH 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-620-694A-10

Query Match 4.8%; Score 176; DB 2; Length 866;

Best Local Similarity 21.8%; Pred. No. 8.3e-08;

Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

QY 237 KNLITLNGPQITLNLN-----HT---DLVPLCLCTQVWPLEPDS-----VRTN-- 271
Db 66 RNLTPSPKDLQQLHFAHTQQGDLFPVAHIE-WTLQTDASILYLEGAELSVLQNTNER 124

QY 272 IC---PFREDPRAHQLWQAARLLTLTLOSLLDAPCS-----LPAEALCWRAPCGD 321
Db 125 LCVREFLEFSLKRRHRRWR-----FTSFHVDPDQYEVYVHLFPK-----IPDGD 172
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSEKLOLOECLWADSLGPKL 360
Db 173 PNHQSNFLVDPCEHARMKVTPCMSSGSLWDPNITVETLEAHQLRVSTFLWNES----- 227
QY 361 DVLLELTRGPQDNRSICALPESGCTS---LPSKASTRAARLGEY-----LLQDLQ 408
Db 228 THYQILLTSPH-----MENHSCFEHMHIP-----APRPEEFHQRSNVTLTLRNK 274
QY 409 S-----GQCLQLWDDDL--GALWACP-----MDKYTHKRWALWLA 442
Db 275 GCCRHQVQIQPFSSCL---NDCLRHSATVSCPEMPDTPPIPDYMLWY--WFITGIS 329
QY 443 CLLFAAALSILILLKKDHAKGWLRLKQDVR--SGAAA-----RGRAALLYSADDS 492
Db 330 ILLVGSVILLIVCMTWRLAGPGSEKYSDDTKYTDGLPAADLIPPLKPRKVIISADHP 389
QY 493 GFERLVGALAS-ALCQLPLRVAVDILWSRRELSAQGPVAVMFHAQRROTLOEGGVVVLIFSP 551
Db 390 LYVDVVVKFAQFLLTACGTEVALDLLEEQAISEAGVMTWVGRKQKQEWVESNSKIIVLCR 449
QY 552 GAVALCSEWLQDGVSGPGA-----HGP--HDAFRASLSCVLPDFLQGRAPGSYVACF 602
Db 450 GTRA---KW--QALLGRGAPVRLRCDHGKPGVGLFTAAAMNMLPDKRPACFGTYVVCYF 504
QY 603 DRLLHPDAVPALEFRTVPVFTLPSQLPDLFLGALQOQPRAPSRGRLOERAEQVSRALQPALDS 662
Db 505 SEVSGDGVDPDLFGAAPRYPLMDRFEVYFRIODLEMFOFGRH-HRVGELSG-----DN 557
QY 663 YEHPPG 668
Db 558 YLRSPG 563
RESULT 3
US-09-022-255-10
; Sequence 10, Application US/09022255
; Patent No. 6072033
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,255
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/620,694
; FILING DATE: 21 MARCH 1996
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/410,535
; FILING DATE: 23 MARCH 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 866 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
us-09-022-255-10

Query Match 4.8%; Score 176; DB 3; Length 866;
Best Local Similarity 21.8%; Pred. No. 8.3e-08;
Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

QY 237 KNL--GPOITLN---HT---DLVPCLCIOVWPLEPDS-----VRTN-- 271
DB 66 RNLTPSPKDLQIOLHFAHTQOGLFPVAHIE-WLQTDASILYLEGALSVLQNTNER 124
QY 272 IC---PFREDPRAHONLQWAAARLLTLQSWLLDAPCS-----LPAAALCWRAPGG 321
DB 125 LCVREFEFLSKLRHHRWR-----FTFSHFVDPDQEVYVHLPKP-----IPDGD 172
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSEKLOECLWADSLGPK 360
DB 173 PNHQSKNFLVPDCEHARMKVTPCMSSGSLWDPNITVETLEAHQLRVFTLWNES----- 227
QY 361 DDVLLLETRGPQDNRSCLALEPSGCTS-----LPSKASTRAARLGEY-----LLQDLQ 408
DB 228 THYQILLTSPH-----MENHSCFEHMHIP-----APRPEEFHORSNVTTLRNK 274
QY 409 S-----GQCQLQWDDDL--GALWACP-----MDKYTHKRWALVWLA 442
DB 275 GCRHVOIQPFSSCL---NDCLRHSAIVSCPEMDTPEPIDYMLVWY--WFTGIS 329
QY 443 CLLFAAALSILLKKDHAKGWLRLKQDVR--SGAAA-----RGRAALLYSADDS 492
DB 330 ILLVGSVILLIVCMTWRLAGPSEKYSDDTKYTDGLPAADLIPPLKPKRWIISADHP 389
QY 493 GFERLVGALAS-ALCOLPLRVAVDLWSRRELSAOGPVANFHAORROTLOEGGVVLLFSP 551
DB 390 LYVDVWLKFAQLLTACGTEVALDLLEEQAISEAGVMTWVGROKQEMVESNKIIVLCR 449
QY 552 GAVALCSEWLQDVGSGPGA-----HGP--HDAFRASLSCVLPDFLQGRAPGSYVACF 602
DB 450 GTRA---KW--QALLGRGAPVRLCDHGKPVGDGLFTAAMNMLLPDKRPACFTYVVCYF 504
QY 603 DRLLHPDAVPALFRTVPVFTLPSQLPDFLQALQOPRAPRSGRLQERAEQVSRALQPALDS 662
DB 505 SEVSCDGVDPDLFGAAPRYPLMDRPEEVYFRIQDLEMFQGRM-HRVGELSG-----DN 557
QY 663 YHPPG 668
DB 558 YLRSPG 563

RESULT 4
US-09-022-696-10
Sequence 10, Application US/09022696
Patent No. 6072037
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle

STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA: US/09/022,696
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 866 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
us-09-022-696-10

Query Match 4.8%; Score 176; DB 3; Length 866;
Best Local Similarity 21.8%; Pred. No. 8.3e-08;
Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

QY 237 KNL--GPOITLN---HT---DLVPCLCIOVWPLEPDS-----VRTN-- 271
DB 66 RNLTPSPKDLQIOLHFAHTQOGLFPVAHIE-WLQTDASILYLEGALSVLQNTNER 124
QY 272 IC---PFREDPRAHONLQWAAARLLTLQSWLLDAPCS-----LPAAALCWRAPGG 321
DB 125 LCVREFEFLSKLRHHRWR-----FTFSHFVDPDQEVYVHLPKP-----IPDGD 172
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSEKLOECLWADSLGPK 360
DB 173 PNHQSKNFLVPDCEHARMKVTPCMSSGSLWDPNITVETLEAHQLRVFTLWNES----- 227
QY 361 DDVLLLETRGPQDNRSCLALEPSGCTS-----LPSKASTRAARLGEY-----LLQDLQ 408
DB 228 THYQILLTSPH-----MENHSCFEHMHIP-----APRPEEFHORSNVTTLRNK 274
QY 409 S-----GQCQLQWDDDL--GALWACP-----MDKYTHKRWALVWLA 442
DB 275 GCRHVOIQPFSSCL---NDCLRHSAIVSCPEMDTPEPIDYMLVWY--WFTGIS 329
QY 443 CLLFAAALSILLKKDHAKGWLRLKQDVR--SGAAA-----RGRAALLYSADDS 492
DB 330 ILLVGSVILLIVCMTWRLAGPSEKYSDDTKYTDGLPAADLIPPLKPKRWIISADHP 389
QY 493 GFERLVGALAS-ALCOLPLRVAVDLWSRRELSAOGPVANFHAORROTLOEGGVVLLFSP 551
DB 390 LYVDVWLKFAQLLTACGTEVALDLLEEQAISEAGVMTWVGROKQEMVESNKIIVLCR 449
QY 552 GAVALCSEWLQDVGSGPGA-----HGP--HDAFRASLSCVLPDFLQGRAPGSYVACF 602
DB 450 GTRA---KW--QALLGRGAPVRLCDHGKPVGDGLFTAAMNMLLPDKRPACFTYVVCYF 504
QY 603 DRLLHPDAVPALFRTVPVFTLPSQLPDFLQALQOPRAPRSGRLQERAEQVSRALQPALDS 662

Db 505 SEVSCDGDVDFLGAAPRIPLMDRFEVFIQDLEMFQGRM-HRVGELSG-----DN 557
QY 663 YFHPPG 668
Db 558 YLRSPG 563

RESULT 5
US-08-978-773-4
; Sequence 4, Application US/08978773
; Patent No. 6083906
; GENERAL INFORMATION:
; APPLICANT: Troutt, Anthony
; TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple PowerMacintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,773
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/052,525
; FILING DATE: 27 NOVEMBER 1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2623-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-978-773-4

Query Match 4.8%; Score 176; DB 3; Length 866;
Best Local Similarity 21.8%; Pred. No. 8.3e-08;
Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

QY 237 KNLTL--GPOIITLN-----HT--DLVPCLCIQVWLEPDS-----VRTN-- 271
Db 66 RNLTPSSPKDLQIQLHFAHQGDLPVVAHIE-WLQTDASILYLEGAEISVLQNTNER 124
QY 272 IC---PFREDPRAHQLWQAARLLRLTQSLWLLDAPCS-----LPAAALCWRAPEGD 321
Db 125 LCVRFEEFLSKLRHHRRWR-----FTFSHFVVDPOQEVTVHHLKP-----IPGD 172
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSKQLQCELMWADSLGPKL 360
Db 173 PNHQSKNLFVPCDEHARKMKTTPCMSSGSLWDPNITVETLEAHQLRVSEFTLWNES----- 227
QY 361 DDVLLLETGPGQDNRLCALPESGCTS-----LPSKASTRAARLGEY-----LLODLQ 408
Db 228 THYQILLTSFPH-----MENHSCFEMHHIP-----APRPEEFHQRSNVTILRLNK 274
QY 409 S-----GQCLQLWDDDL--GALWACP-----MDKYIHKRWALVWLA 442

Db 275 GCCRHQVOIQFFSSCL---NDCLRHSATVSCPEMPDTPPIPDYMPLMWVY--WFIITGIS 329
QY 443 CLLEFAAALSILLKDKHAKGWLRLKKODVR--SGAAA-----RCRAALLLYSADDS 492
Db 330 ILLVGSVILLIVCTWRLAGPGSEKYSDDTKYTDTGLPAADLLPPPLKPKRWIYISADHP 389
QY 493 GFERLVGALAS-ALCQLPLRVAVDLMSRRLESAQGPVAMFHAORRQTLQEGGVVLLFSP 551
Db 390 LYVDVLKFAQFLLTAGCTEVALDLEBOAISEAGVMTWGRQKQEMVESKSIIVLCR 449
QY 552 GAVALCSEWLDGVSGPGA-----HCP--HDAFRASLSCLVLPDFLOGRAGPSYVGCACF 602
Db 450 GTRA---KW--QALLGRGAPVRLRCDHGKPVGDLFTAMNMILPDFKPKPACFGTYVWCYF 504
QY 603 DRLLHPDPAVPAFRTVPVFTLPSLPDPLGALQOPRAPRPSRGLQERAEQVSRALQPALDS 662
Db 505 SEVSCDGDVDFLGAAPRIPLMDRFEVFIQDLEMFQGRM-HRVGELSG-----DN 557
QY 663 YFHPPG 668
Db 558 YLRSPG 563

RESULT 6
US-09-022-253-10
; Sequence 10, Application US/09022253
; Patent No. 6096305
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6096305e1 Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,253
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694
; FILING DATE: 21-MARCH-1996
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/410,535
; FILING DATE: 23 MARCH 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-022-253-10

Query Match 4.8%; Score 176; DB 3; Length 866;

Best Local Similarity 21.8%; Pred. No. 8.3e-08;

Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

QY 237 KNL--GPOIITLN---HT---DLVPCICIQVWPLEPDS-----VRN-- 271
DB 66 RNLTPSSPKDLQQLHFAHTQOQDLFPVAHIE-WTLOTDA SILYLEGAELSVLQNTN 124
QY 272 IC---PFREDPRAHONLWQAARLLTLQSWLLDAPCS-----LPAEALCWRAPG 321
DB 125 LCVRFEEFLSKLHHRRWR-----FTFSHFVDPDOEYEVTVHHLKP-----IPGD 172
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSEKIQCEIWLADSLGPK 360
DB 173 PNHQSKNPLVPCDCEHARKVTPCMSSGSLWDPNITVETLEAHQLRVSTLWNES----- 227
QY 361 DDVLLLETRGPQDNRLCALPESGCTS-----LPSKASTRAARLGEY-----LLQDLQ 408
DB 228 THYQILLTSFPH-----MENHSCFEHMHIP-----APRPEEFHORSNVTLTLRNLK 274
QY 409 S-----GQCQLQWDDDL--GALWACP-----MDKYIHKRWALVWLA 442
DB 275 GCRHVOVQIOFFSSCL---NDCLRHSATVSCPEMPDTPPIPDYMLWVY--WFTGIS 329
QY 443 CLLFAAALSLILLKKDHAKGWLRLLLKQDVR--SGAAA-----RGRAALLYSADDS 492
DB 330 ILLVGSVILLIVCMWRLAGPGSEKYSDDTKYTDGLPAADLIPPLPKPKRWIYISADHP 389
QY 493 GFERLVGALAS-ALCOLPLRVAVDLWSRRELSAQGPVAFWFAHQRROTLOEGGVVLLFSP 551
DB 390 LYVDVWLKFAQFLTTACGTEVALDLLEQAISEAGVMTWVGROKQEMVESNSKIIVLCR 449
QY 552 GAVALCSEWLQDGVSGPGA-----HGP--HDAFRASLSCLVLPDFLOGRAPGSYVACF 602
DB 450 GTRA---KW--QALLGRGAPVRLCDHGKPVGDLFTAAMNMLPDKFRKPCFGYVVCYF 504
QY 603 DRLLHPDAVPALFRTVPVFTLPSQLPDFLGAQQPRAPRSGRGLQERAEQVSRALQPALDS 662
DB 505 SEVSCDGDVDFLEGAAPRYPLMDRFEVYFRIQDLEWFPQGRM-HRVGELSG-----DN 557
QY 663 YFHPGP 668
DB 558 YLRSPG 563

RESULT 7

US-09-022-260-10

; Sequence 10, Application US/09022260

; Patent No. 6100235

; GENERAL INFORMATION:

; APPLICANT: Yao, Zhengbin

; APPLICANT: Springs, Melanie

; APPLICANT: Fanslow, William

; TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Immunex Corporation

; STREET: 51 University Street

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: Apple Power Macintosh

; OPERATING SYSTEM: Apple Operating System 7.5.5

; SOFTWARE: Microsoft Word for Apple, Version 6.0.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/022,260

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION NUMBER: 08/620,694

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: USSN 08/410,535

; FILING DATE: 23 MARCH 1995

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Perkins, Patricia Anne

; REGISTRATION NUMBER: 34,695

; REFERENCE/DOCKET NUMBER: 2617-B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206)587-0430

; TELEFAX: (206)

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 866 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-022-260-10

Query Match

4.8%; Score 176; DB 3; Length 866;

Best Local Similarity 21.8%; Pred. No. 8.3e-08;

Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

QY 237 KNL--GPOIITLN---HT---DLVPCICIQVWPLEPDS-----VRN-- 271
DB 66 RNLTPSSPKDLQQLHFAHTQOQDLFPVAHIE-WTLOTDA SILYLEGAELSVLQNTN 124
QY 272 IC---PFREDPRAHONLWQAARLLTLQSWLLDAPCS-----LPAEALCWRAPG 321
DB 125 LCVRFEEFLSKLHHRRWR-----FTFSHFVDPDOEYEVTVHHLKP-----IPGD 172
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSEKIQCEIWLADSLGPK 360
DB 173 PNHQSKNPLVPCDCEHARKVTPCMSSGSLWDPNITVETLEAHQLRVSTLWNES----- 227
QY 361 DDVLLLETRGPQDNRLCALPESGCTS-----LPSKASTRAARLGEY-----LLQDLQ 408
DB 228 THYQILLTSFPH-----MENHSCFEHMHIP-----APRPEEFHORSNVTLTLRNLK 274
QY 409 S-----GQCQLQWDDDL--GALWACP-----MDKYIHKRWALVWLA 442
DB 275 GCRHVOVQIOFFSSCL---NDCLRHSATVSCPEMPDTPPIPDYMLWVY--WFTGIS 329
QY 443 CLLFAAALSLILLKKDHAKGWLRLLLKQDVR--SGAAA-----RGRAALLYSADDS 492
DB 330 ILLVGSVILLIVCMWRLAGPGSEKYSDDTKYTDGLPAADLIPPLPKPKRWIYISADHP 389
QY 493 GFERLVGALAS-ALCOLPLRVAVDLWSRRELSAQGPVAFWFAHQRROTLOEGGVVLLFSP 551
DB 390 LYVDVWLKFAQFLTTACGTEVALDLLEQAISEAGVMTWVGROKQEMVESNSKIIVLCR 449
QY 552 GAVALCSEWLQDGVSGPGA-----HGP--HDAFRASLSCLVLPDFLOGRAPGSYVACF 602
DB 450 GTRA---KW--QALLGRGAPVRLCDHGKPVGDLFTAAMNMLPDKFRKPCFGYVVCYF 504
QY 603 DRLLHPDAVPALFRTVPVFTLPSQLPDFLGAQQPRAPRSGRGLQERAEQVSRALQPALDS 662
DB 505 SEVSCDGDVDFLEGAAPRYPLMDRFEVYFRIQDLEWFPQGRM-HRVGELSG-----DN 557
QY 663 YFHPGP 668
DB 558 YLRSPG 563

RESULT 8

US-09-022-259-10

; Sequence 10, Application US/09022259

QY 237 KNT--GPOIITL---HT---DLVPCLCIQVWPLEPDS-----VRN-- 271
Db 66 RNLTPSSPKDLQLHFAHTQOGDLFPVAHIE-WLQTDASILYLEGAELSVQLNTNR 124
QY 272 IC---PFREDPRAHQLWQAARLLTLQSWLLDAPCS-----LPAEALCWRAPGD 321
Db 125 LCVREFELSKLRHHRRWR-----FTFSHFVYDPOEYEVTVHHLKP-----IPGD 172
QY 322 PCOP-----LVPPLS-----WE-NVTVDVNSSEKLOEQCLWADSLGPK 360
Db 173 PNHQSKNFLVPDCEHARMKVTPCMSSGSLWDPNITVETLEAHQLRVSTLWNES----- 227
QY 361 DDVLLLETRGPQDNRSICALEPSGCTS-----LPSKASTRAARLGEY-----LLQDLQ 408
Db 228 THYQILLTSPH-----MENHSCFEHMHIP-----APRPEEFHQRNSVTLTLENK 274
QY 409 S-----GOCQLQWDDDL---GALWACP-----MDKYTHKRWALWLA 442
Db 275 GCRHQVOIQPFSSCL---NCLRHSAIVSCPEMPDTPPIPDYMLVY--WFTGIS 329
QY 443 CILFAAALSILLLKKDHAKWMLRLKQDVR---SGAAA-----RGAALLLSYSDS 492
Db 330 ILLVGSVILLVCMTRWLAGPGESEKSDTKYTDGLPAADLIPPLPKPRKVIISADHP 389
QY 493 GFERLVGALAS-ALCOLPLRVAVDLWSRRELSAQGPVAFHQAQRQTLOEGGVVLLFSP 551
Db 390 LVYDVVVKFAQFLLTACGTGEVALDLLEQAISEAGVMTWVGROKQEVESNKIIVLCR 449
QY 552 GAVALCSEWLQGVSGPA-----HGP--HDAFRASLSCLVLPDFLQGRAPSYVGACF 602
Db 450 GTRA---KW--QALLGRGAPVRLRCDHGXPKVGLDFTAAAMWMLPDKRACFTGVVYCYF 504
QY 603 DRLLHPDPAVPLFRVTVPLTSQLPDLFGALQOQPRAPRSGRLOERAEQVSRALQPADS 662
Db 505 SEVSCGDGVDPLFGAAPRYPLMDRFEFVYFRIQDLEMFQGRM-HRVGELSG-----DN 557
QY 663 YFHPG 668
Db 558 YLRSPG 563

RESULT 10

US-08-620-694A-2
; Sequence 2, Application US/08620694A
; Patent No. 5869286

GENERAL INFORMATION:

; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694A
; FILING DATE: 21 MARCH 1996
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/410,535

; FILING DATE: 23 MARCH 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 864 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-620-694A-2

Query Match 4.58; Score 166; DB 2; Length 864;
Best Local Similarity 20.4%; Pred. No. 7.4e-07;
Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;

QY 192 PALPW---LNVASADGNVHLVLN---VSEEOHFGSLYWNVOGPP--KPRW--HKNL 239
Db 14 PALGWLILLNLVAPGRASPRLLDFPAPVCAQE--GLSC---RVKNSTCLDDSIHFNKL 68
QY 240 T--GPOIITL-----HTDLVPCLCIQVWPLEPDS-----VRN--IC- 273
Db 69 TPSPKNIYINLSVSTQHGLVPLVHVE-WLQTDASILYLEGAELSVQLNTNERLCV 127
QY 274 --PREDPRAHQLWQAARLLTLQSWLLDAPCSLPAE-----ALCWAPAGDP--- 322
Db 128 KFQFLSMLQHRKRWR-----FSFSHFVVD---PQOEYEVTVHHLPKPIPDGPNHK 176
QY 323 -----COPLVPPLS-----WE-NVTVDVNSSEKLOEQCLWADSLGPKDDVL 364
Db 177 SKIIFVPCDESKMKMTTSCVSSGSLWDPNITVETLQHLRVFTLWNEST-PYQ--V 232
QY 365 LLETGRPDNRSICALEPSGCTSL-----PSKASTRAARLGEYLLQDLSGOCQLQWDD 418
Db 233 LLESFSDSENHS-----CFDVVKQIFAPRQEEFHQRANVTFTLSKFH----- 274
QY 419 DLGALWAC-----PMDKYIHKRW-----LV 439
Db 275 ---WCCHHHVQVQPFSSCLNCLRHAVTVPCPVISNTTVPKPVADYI-PLWVYGLIT 328
QY 440 WLACLLFAAALSILLLKKDHAKWMLRLKQDVRSGA-----SGDOEKHGGDSKINGILPVADLTTP 476
Db 329 LIALLVGSVIVLTCMT-----W-RL-----SGDOEKHGGDSKINGILPVADLTTP 375
QY 477 AARGAALLLYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSAQGPVAFHQAQR 535
Db 376 PLRPRKVIIVYSADHPLYVEVVLKFAQFLITACGTGEVALDLLEEQVISEVGVMTWVSQK 435
QY 536 RQTLOEGGVVLLFSPG-----AVAL-CSEVLDQGVSGPAGHGDHAFAS 580
Db 436 QEMVESNKIILCSRGTOAKWKAILGWAEPAYQLRCDHMKPAG-----DLFTAA 485
QY 581 LSCVLPDFLOGRAPGSYVGACFDRLLHPDPAVPLFRVTVPLTSQLPDLFGALQOQPRAP 640
Db 486 MNMLPDKRACFTGVVYFSGICSERDVPDLNFTSRYPLMDRFEFVYFRIQDLEMF 545
QY 641 RSGRLOERAEQVSRALQPADLSYFHPH 667
Db 546 EPGRMHHVRELGTG-----DNYLQSP 565

RESULT 11

US-09-022-255-2
; Sequence 2, Application US/09022255
; Patent No. 6072033
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin

APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,255
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/620,694
FILING DATE: 21 MARCH 1996
APPLICATION NUMBER: USSN 08/538,765
FILING DATE: 7 AUGUST 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-255-2

Query Match 4.5%; Score 166; DB 3; Length 864;
Best Local Similarity 20.4%; Pred. No. 7.4e-07;
Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;

QY 192 PALPW-----LNVSADGDNVHLVILN-----VSEQHFGSLYWNQVQGP--KPRW--HKNL 239
Db 14 PALGWLLLLLNLVLAAGRASPRLLDFPAPVCAQE--GLSC---RVKNSTCLDDSWIHPKNL 68
QY 240 T--GPOIITN-----HTDLVPLCIQWPLEPDS-----VRTN--IC- 273
Db 69 TPSSPKNIYINLSVSTQHGELVPVHLVE-WTLQTDASILYLEGAEISVLQNTNERLCV 127
QY 274 --PFREDPRAHQLWQAARLLTLQSLWLDAPCSLPAAE-----ALCWRAPOGDP--- 322
Db 128 KQQLSLMLQHHRRWR-----FSESHFVVD-----PGQYEYTVVHLPKPIPGDGNHK 176
QY 323 -----COPLVPPPLS-----WE-NVTVDVNSSEKLQLOECLWADSLGPKLDDVL 364
Db 177 SKIIFVPDCEDSKMKMTTSCVSSGSLWDPNITVETLDTQHLRVDFTLWNEST-PYQ---V 232
QY 365 LLETGCPQDNRSICALPEGSCTSL-----PSKASTFAARLGEYLLQDLOGGQCLQLWDP 418
Db 233 LLESFSDSENHS-----CFDVVKQIFAPROEHEFHQRANVTFTLSKFH----- 274
QY 419 DLGALWAC-----PMDKYIHKRW-----LV 439
Db 275 -----WCCHHHVQVQFFSSCLNDCLRHAVTVPCVISNTVTPKPVADYI-PLWVYGLIT 328
QY 440 WLACLLFAAALSILLKKDKHAKGWLRLKQDVRSGA----- 476

Db 329 LIALLLGVSVILICMT-----W-RL-----SGAQEKHGDDSKINGILPVADLTPP 375
QY 477 AARGAALLLYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSSAQGPFVAMFHAQR 535
Db 376 PLRPRKWIVIVSADHPLVVEVVLKFAQLITACGTEVALDLLEEQVISEVGVMTWVSQK 435
QY 536 RQTLOEGGVVLLFSPG-----AVAL-CSENLQDVGSGGANGHPHDAFRAS 580
Db 436 QEMVESNSKIILCSRGTOAKWAKILGWAEPVQLRCDHWKPAG-----DLFTAA 485
QY 581 LSCVLPDFLOGRAPGSGYVACFDRLHLLHDVAPALFRTVPVTLPSQLPDFLGCALQOPRAP 640
Db 486 MNMILPDFKRPACFGTYVYVCFSGICSRDVPDLFNITSRYPLMDRFEVVFRIODLEMF 545
QY 641 RSGRLQERAEQVSRALQALPDSYFHP 667
Db 546 EPGRMHVRELTG-----DNYLQSP 565
RESULT 12
US-09-022-696-2
Sequence 2, Application US/09022696
Patent No. 6072037
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,696
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-696-2

Query Match 4.5%; Score 166; DB 3; Length 864;
Best Local Similarity 20.4%; Pred. No. 7.4e-07;
Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;

QY 192 PALPW-----LNVSADGDNVHLVNLN-----VSEQHFGSLYWNQVQPP--KPRW--HKNL 239
DB 14 PALGWLILLNLNVLAPGRASPRLLDFFPAPVCAQE--GLSC---RVKNSTCLDDSWIHPKNL 68
QY 240 T--GPOIITLN-----HTDLVPCLCIQVWPLEPDS-----VRTN--IC- 273
DB 69 TPSSPKNIYINLSVSTQHGLVPLVHVE-WTLOTDA SILYLEGAELSVLQNTNERLCV 127
QY 274 ---PFREDPRAHONLWOAARLLTLQSWLLDAPCSLPAAE-----ALCWRAPGDDP--- 322
DB 128 KFQFLSMLQHRKRW-----FSFHFVVD-----PGQYEVTVHHLPRIPDGDPNHK 176
QY 323 -----COPLVPPLS-----WE-NVTVDVNSSEKQLQECLEWADSLGPKLDDVL 364
DB 177 SKIIFVPDCEDSKMKTTCVSSGSLWDPNITVETLDTQHLRVDFTLWNEST-PYQ---V 232
QY 365 LLETGPGDNRSICALPESGCTSL-----PSKASTRAARLGEYLQDLQSGCQLQWLDD 418
DB 233 LLESTSDSENHS-----CFDVVKQIFAPRQEEFHQRANVTFTLSKFH----- 274
QY 419 DLGALWAC-----PMDKYIHKRWA-----LV 439
DB 275 -----WCCHHHVQVOPFFSSCLNCLRHAVTPCPVISNTTVPKPVADYI-PLWVYGLIT 328
QY 440 WLACILFAAALSILLKLDHAKGWLRLKQDVRSGA----- 476
DB 329 LIAILLVGSVILLICMT-----W-RL-----SGAQEKHGDDSKINGILPVADITPP 375
QY 477 AARGRAALLYSADDSGFERLVGALASAL--COLPLRVAVDLWSRRELSAOGPVAFHAQR 535
DB 376 PLRPRKVMIVYSADHPLYVEVVLKFAQLITACGTEVALDLLEEQVISEVGVMTWVSROK 435
QY 536 RQTLQEGGVVLLFSPG-----AVAL-CSEWLODGVSGGAGHPHDAFRAS 580
DB 436 QEMVESNSKIIILCSRGTOAKWKAILGWAEPAVOLRCDHMKPAG-----DLFTAA 485
QY 581 LSCVLPDFLOGRAGSYGACFDRLLHPDAVPALFRTVPVFTLPSQLPDFLGAQQPRAP 640
DB 486 MNMILPDFKRPACFGTYVYVYFSGICSERDVPDLFNITSRYPLMDRFEVYFRIOQLEMF 545
QY 641 RSGRLQERAEQVSRALQALPDLSYFHP 667
DB 546 EPRGHHVRELGT-----DNYLQSP 565

RESULT 13

US-08-978-773-2
; Sequence 2, Application US/08978773
; Patent No. 6083906
; GENERAL INFORMATION:
; APPLICANT: Troutt, Anthony
; TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple PowerMacintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,773
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/052,525
; FILING DATE: 27 NOVEMBER 1996

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2623-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-978-773-2

Query Match 4.5%; Score 166; DB 3; Length 864;

Best Local Similarity 20.4%; Pred. No. 7.4e-07;
Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;

QY 192 PALPW-----LNVSADGDNVHLVNLN-----VSEQHFGSLYWNQVQPP--KPRW--HKNL 239
DB 14 PALGWLILLNLNVLAPGRASPRLLDFFPAPVCAQE--GLSC---RVKNSTCLDDSWIHPKNL 68
QY 240 T--GPOIITLN-----HTDLVPCLCIQVWPLEPDS-----VRTN--IC- 273
DB 69 TPSSPKNIYINLSVSTQHGLVPLVHVE-WTLOTDA SILYLEGAELSVLQNTNERLCV 127
QY 274 ---PFREDPRAHONLWOAARLLTLQSWLLDAPCSLPAAE-----ALCWRAPGDDP--- 322
DB 128 KFQFLSMLQHRKRW-----FSFHFVVD-----PGQYEVTVHHLPRIPDGDPNHK 176
QY 323 -----COPLVPPLS-----WE-NVTVDVNSSEKQLQECLEWADSLGPKLDDVL 364
DB 177 SKIIFVPDCEDSKMKTTCVSSGSLWDPNITVETLDTQHLRVDFTLWNEST-PYQ---V 232
QY 365 LLETGPGDNRSICALPESGCTSL-----PSKASTRAARLGEYLQDLQSGCQLQWLDD 418
DB 233 LLESTSDSENHS-----CFDVVKQIFAPRQEEFHQRANVTFTLSKFH----- 274
QY 419 DLGALWAC-----PMDKYIHKRWA-----LV 439
DB 275 -----WCCHHHVQVOPFFSSCLNCLRHAVTPCPVISNTTVPKPVADYI-PLWVYGLIT 328
QY 440 WLACILFAAALSILLKLDHAKGWLRLKQDVRSGA----- 476
DB 329 LIAILLVGSVILLICMT-----W-RL-----SGAQEKHGDDSKINGILPVADITPP 375
QY 477 AARGRAALLYSADDSGFERLVGALASAL--COLPLRVAVDLWSRRELSAOGPVAFHAQR 535
DB 376 PLRPRKVMIVYSADHPLYVEVVLKFAQLITACGTEVALDLLEEQVISEVGVMTWVSROK 435
QY 536 RQTLQEGGVVLLFSPG-----AVAL-CSEWLODGVSGGAGHPHDAFRAS 580
DB 436 QEMVESNSKIIILCSRGTOAKWKAILGWAEPAVOLRCDHMKPAG-----DLFTAA 485
QY 581 LSCVLPDFLOGRAGSYGACFDRLLHPDAVPALFRTVPVFTLPSQLPDFLGAQQPRAP 640
DB 486 MNMILPDFKRPACFGTYVYVYFSGICSERDVPDLFNITSRYPLMDRFEVYFRIOQLEMF 545
QY 641 RSGRLQERAEQVSRALQALPDLSYFHP 667
DB 546 EPRGHHVRELGT-----DNYLQSP 565

RESULT 14

US-09-022-253-2
; Sequence 2, Application US/09022253
; Patent No. 6096305
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie

Query Match 4.5%; Score 166; DB 3; Length 864;
Best Local Similarity 20.4%; Pred. No. 7.4e-07;

```
Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;
QY 192 PALPW-----LNVSADGDNVHLN-----VSEQHFGLSLYWNOVQGP--KPRW--HKNL 239
Db 14 PALGWLILLNLNVLAPGRASRLDFAVPVCAQE--GLSC---RVKNSTCLDDSDWIHPKNL 68
QY 240 T--GPOIITLN-----HTDLVPCICIQVWPLEPDS-----V--VRTN--IC- 273
Db 69 TPSSPKNIYINLSVSTQHGELVPVLHVE-WTLOTDASILYLEGAELSVLQLTNTERLCV 127
QY 274 --PFREDPRAHQNLAARLRLTLQSWLWDAPCSLPAEA-----ALCWRAPGGDP--- 322
Db 128 KFOFLSMLQHHRKRWR-----FSFSHFVVD-----PGQEVETVHHLPKPIPDGDPNHK 176
QY 323 -----COPLVPPLS-----WE-NVTVDVNSSEKLQLOECLWADSLGPKDDVL 364
Db 177 SKIIFVPDCEDSKMKTTCVSSGSLMDPNITVETLDTQHLRVDFTLWNEST-PYQ---V 232
QY 365 LLETGPGQDNRSICALEPSCGCTSL-----PSKASTRAARLGEVLLQDLQSGQCQLWDD 418
Db 233 LLESFSDSENHS-----CFDVVKQIFAPROEEFHQRANVTFTLSKFH----- 274
QY 419 DLGALWAC-----PMDKYIHKRWA---LV 439
Db 275 -----WCCHHHVQVQPFSSCLNDCLRHAHTVPCPVISNTTVPKPVADYI-PLWVYGLIT 328
QY 440 WLACLLFAAALSILLKXKHAKGWLRLKQDVRSGA----- 476
Db 329 LIAILLVGSVIVLIICMT-----W-RL-----SGADQEKHGDDSKINGILPVADLTTP 375
QY 477 AARGRAALLYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSAQGPVAFHAQR 535
Db 376 PLRPRKVIWIVYSADHPLYVEWLKFAQLITACGTEVALDLEEQVISEVGMVTWVSQK 435
QY 536 ROTLOEGGVVVLGFSFG-----AVAL-CSEWLQDGVSGGAGHPHDAFRAS 580
Db 436 QEWVESNSKIIILCSRGTOAKWAILGWAEPVQLRCDHWKPG-----DLFTAA 485
QY 581 LSCVLPDFLQGRAPGSYGACFDRLHPDAPALFRTVPVFTLPFSQLPDLFALQOQPRAP 640
Db 486 MNMILPDFKRPACFGTYVVCYFSGICSERDVPDLFNITSRYPLMDRFEVYFRIQDLEMF 545
QY 641 RSGRLQERAEQVSRALQPALDSTFHPP 667
Db 546 EPGRMHVRELGTG-----DNYLQSP 565
```

Search completed: May 19, 2002, 14:10:18
Job time: 6375 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 12:24:52 ; Search time 63.07 seconds
(without alignments)

1054.286 Million cell updates/sec

Title: US-09-608-918-2

Perfect score: 3716

Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPGRGVGPGAGDGT 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	137.5	3.7	3707	2 S18252	heparan sulfate pr
2	130.5	3.5	1366	2 T35985	probable large pro
3	115.5	3.1	745	2 H85840	hypothetical prote
4	114.5	3.1	745	2 C99995	hypothetical prote
5	114.5	3.1	1193	2 F83264	hypothetical prote
6	113.5	3.1	761	2 S20458	pqqf protein - kle
7	111.5	3.0	478	2 D75564	hypothetical prote
8	108	2.9	2629	2 T30987	telomerase-associ
9	107	2.9	719	2 T35189	probable ATP-depen
10	107	2.9	825	1 A60386	interleukin-4 rece
11	107	2.9	1118	1 A49724	protein-tyrosine-p
12	104.5	2.8	238	2 C70841	hypothetical prote
13	104.5	2.8	975	2 S33121	homeotic protein C
14	104.5	2.8	1711	1 A55148	protein-tyrosine-p
15	104	2.8	1487	2 T02850	hypothetical prote
16	103.5	2.8	676	1 WMBEX6	ATP-dependent heli
17	103.5	2.8	828	2 AD0412	hypothetical prote
18	103.5	2.8	2055	2 T00093	hypothetical prote
19	103	2.8	4845	2 T31067	BIR repeat contain
20	101.5	2.7	1151	2 S48431	probable membrane
21	101	2.7	274	2 A83583	probable biotin sy
22	100.5	2.7	709	2 F75584	hypothetical prote
23	100.5	2.7	741	2 T31164	hypothetical prote
24	100.5	2.7	753	1 W2BPE8	gene 56 protein -
25	99	2.7	429	2 T36088	probable secreted
26	99	2.7	636	2 B83513	probable heat choc
27	99	2.7	660	2 E98169	serine proteinase
28	99	2.7	660	2 AB3118	serine proteinase
29	99	2.7	679	2 B75262	conserved hypothet

Query Match 3.7% ; Score 137.5; DB 2; Length 3707;

hypothetical prote
enhancer of split
MEGF8 protein - hu
conserved hypothet
immediate-early pr
155k transcription
nonstructural poly
exo-alpha-sialidas
hypothetical prote
hypothetical prote
erbB kinase activa
polyketide synthas
perlecan precursor
hypothetical prote
kinase-related pro

30 98.5 2.7 724 2 T47149
31 98.5 2.7 741 2 B49555
32 98.5 2.7 1737 2 T00209
33 98 2.6 427 2 E87669
34 98 2.6 1487 1 EDBEE1
35 98 2.6 1487 1 EDBEF6
36 98 2.6 2205 1 MNMVRN
37 97.5 2.6 418 2 JC7588
38 97.5 2.6 617 2 F75484
39 97.5 2.6 698 2 A82593
40 97.5 2.6 850 2 JC5700
41 97.5 2.6 3739 2 T17410
42 97.5 2.6 4391 2 A38096
43 97 2.6 415 2 T46462
44 97 2.6 779 2 AG1978
45 97 2.6 2594 2 A35774

ALIGNMENTS

RESULT 1
S18252
heparan sulfate proteoglycan - mouse
N:Alternate names: perlecan
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C:Accession: S18252; A31917; B31917; S66460
R:Noonan, D.M.; Fullen, A.; Valente, P.; Cai, S.; Horigan, E.; Sasaki, M.; Yamada, Y.;
J. Biol. Chem. 266, 22939-22947, 1991
A:Title: The complete sequence of perlecan, a basement membrane heparan sulfate prote
adhesion molecule.
A:Reference number: S18252; MUID:92078153
A:Accession: S18252
A:Molecule type: mRNA
A:Residues: 1-3707 <NO>
A:Cross-references: EMBL:M77174; NID:920295; PIDN:AAA39911.1; PID:9200296
R:Noonan, D.M.; Horigan, E.A.; Ledbetter, S.R.; Vogeli, G.; Sasaki, M.; Yamada, Y.;
J. Biol. Chem. 263, 16379-16387, 1988
A:Title: Identification of cDNA clones encoding different domains of the basement men
A:Reference number: A92680; MUID:89034110
A:Accession: A31917
A:Molecule type: mRNA
A:Residues: 940-1601 <NO>
A:Cross-references: GB:J04054; NID:9200252; PIDN:AAA39899.1; PID:9200253
A:Accession: B31917
A:Molecule type: mRNA
A:Residues: 1870-2600 <NO3>
A:Cross-references: GB:J04055; NID:9200300; PIDN:AAA39912.1; PID:9200301
R:Schulze, B.; Mann, K.; Battistutta, R.; Wiedemann, H.; Timpl, R.
Eur. J. Biochem. 231, 551-556, 1995
A:Title: Structural properties of recombinant domain III-3 of perlecan containing a g
A:Reference number: S66460; MUID:95377282
A:Accession: S66460
A:Molecule type: protein
A:Residues: 1272-1274, 'X', 1276, 'X', 1278-1279 <SCH>
C:Superfamily: LDL receptor ligand-binding repeat homology: EGF homology; laminin G r
C:Keywords: glycoprotein
F:195-234/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:285-319/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:325-359/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:368-403/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:764-811/Domain: laminin-type EGF-like homology <LEG>
F:1159-1206/Domain: laminin-type EGF-like homology <LEG7>
F:1563-1610/Domain: laminin-type EGF-like homology <EG7>
F:1613-1668/Domain: laminin-type EGF-like homology <LEG8>
F:3163-3198/Domain: EGF homology <EGF>
F:3270-3423/Domain: laminin G repeat homology <EGF>
F:3464-3492/Domain: EGF homology <EGF7>
F:1256,1891,2336,2394,2427/Binding site: carbohydrate (Asn) (covalent) #status predic

Best Local Similarity 19.58; Pred. No. 0.13; Mismatches 235; Indels 365; Gaps 42;
Matches 163; Conservative 73;

```
QY 4 PWFLLSLALGRSPVVLRLVGPQDAT-----HCSP-----GLSC----- 39
Db 272 PQFLPSVPGPS-----ACGPQASCHSGHCIPRDYLCDDGQEDCRDSDGLCCASPP 323
QY 40 -----RLWSDILCLPGDI-----VPAPGVLPATHTLQTEL 71
Db 324 PCBPNEFACENGHCALKLRCD-----GDFCEDRTDEANGSVKQPGVEGCTHFCQVST 378
QY 72 LR-----CQKEDCDCLRLVAVHLAVHGWEEPEDEEKGGAADSGVEPRNASLOAQ 124
Db 379 NRCIPASFHDESDC-----PDRSDEFGCMPPQVVTTPQO-SIOA- 418
QY 125 VLSFQAYPTARCVLLEVQPAALVQFGSVGVVYDCFEAALGSEVRWISYTPRYEKE 184
Db 419 -----SRGQ-----TVTFCTVATGVPTPIINW-----R 441
QY 185 LNHTQOLPALPMLNVSADGNVHLVNVSEOHFG-----LSLYW 224
Db 442 LN-WGHIPAPRPVMTSEGGRTLIIRDVKEADQAVTCEAMNSRGMVFGIPDGVLELV- 499
QY 225 NOVQGPCKPRWHKNLTGQIITLNHTDLPVCLCQVWPLEPDSVRTNCP-----FREDPR 280
Db 500 --PORGPCPDGHFYLE-----DSASCLPCFCFV-----TNVCOSSLRFRDQIR 541
QY 281 A---HONLWQAARLRLTLQSWLLDAPCSLPAEALCWAPGGPCOPVPLPWSWENVTV 337
Db 542 LSFQDQNDKGVNTM-----PSQGPVPLSLSTQLOI 573
QY 338 DKVL-BFPLKGNHLCVQVNSSEKLQLOECLWA-----DSLQ----- 374
Db 574 DPALQEFOL-----VDLSRRFLVHDFAFWALPKQFLGNKVDSYGGFLRYKRYELA 623
QY 375 -----PL-RDDVLL-----LETGRQDNRLCALPESGCTSLPSKASTRAARIGEYLLQ 422
Db 624 RGMLEPVQKPDVILVAGYRLHSGHTPTH-----PGTLNQRQVQLSE----- 666
QY 423 DLQSGOCLQLDWDLDLALWACPMKYIHKRWALVWLACLLFA-AALSILILLKKDHAKAA 481
Db 667 -----EHWVESGRPVQRAEMLOALASLEAVLLQTVYNTKMA 703
QY 482 ARGRAALLYSDSGFERLVGALASAL--COLPLRVAVDLWSRRELSAOGPVAMFAQR 539
Db 704 SVGLSDIVM---DVTVTHTTIHGRAHSVEBCRCPIGVS-----GLSCESCDAHF--- 749
QY 540 RQTLQEGGVVLLFS-----PGAVALC-----SEWLQDGVSGPGAGHGPDAFR 582
Db 750 --TRVEGGPYLGTSCGNCNGHASSCDPVYGHCLNCOHNTGEPQCDKCKPGFFG--DATK 805
QY 583 ASLSCVLPDFLQGRAPGSYGA-----CFDRLHHPDAVPALFRTPVPT---LPSQLP 632
Db 806 ATATACRP-----CFCPTIDASRRSDFYCF---LTDGQATCDACAPGYTGRRCESCAP 856
QY 633 DFILGALOQP-----RAPRSGRLOQRAEQ-----VSRLQPALDYSVFH 669
Db 857 GYEGNFIQPGKCRPTQETVRCDERGSLGTCGTCRCKNNVGVRLCNECSDGSFH 912
```

RESULT 2

T35985 probable large pro/Ala/Gly-rich protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T35985

R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, February 1999

A:Reference number: Z21589

A:Accession: T35985

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1366 <MUR>

A:Cross-references: EMBL:AL035559; PIDN:CAB37473.1; GSPDB:GN00070; SCOEDB:SC9F2.06c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC9F2.06c

Query Match 3.5%; Score 130.5; DB 2: Length 1366;
Best Local Similarity 22.4%; Pred. No. 0.13;
Matches 175; Conservative 57; Mismatches 238; Indels 311; Gaps 42;

```
QY 21 LERLVGPQDATHCISGLSRLWSDILCLPGDIVPAGPVLPATHTLQTELVLRCOKETOC 80
Db 675 LDGALDAPDAGH---PLTIRLLSEVRAALPGP--PAPVEVTRDAVETAYLDL----- 721
QY 81 DLCLRVAVHLAVHGWEEPEDEEKGGA-----DSGVEEPRNASI-QAQVVLSSFOA 131
Db 722 -MCLRVATRLA-----DENGURGTAVRRLAKVSGOVHEAARRSLPGCGGLDRES 771
QY 132 YPT-----ARCVLLE-VQVPAAL-----VQFGQSVG 156
Db 772 FETLPCGPAPARLGGGTGWAPAVLAEGLFVPTGSGYRFAHEELADWIOGTHLDLGEALR 831
QY 157 SVVYDCFEAALGSEVRWISYTOPRYEKELNHTQOLPALPMLNVSADGNVHLVNVSEBQ 216
Db 832 ALVHR-RDTPLGT-----HTRTLP-VPHHRI---GSVVEALLLIA-RQ 868
QY 217 HFGLSLYWNQVQGPCKPRWHKNLTGQIITLNHTDLPVCLCQVWPLEPDSVRTNICPFR 276
Db 869 H-----GVPO-----LALTLEELVHAL-----D 886
QY 277 EDPAHQNLWQAARLRLTLQSWLLDAPCSLPAEAL-----CWAPGGPCOPVLPVP 329
Db 887 RDPHS-----WAAARL---LAELATRPVDPATPYTDVLRLLADGIAERAGDQGTQVQFQ 938
QY 330 LSWENVTVDKVLEFPFLKGNHLCVQVNSSEKLQLOECLWADSLGPKDDVLLLETRGQ 389
Db 939 AFWTAPRVPAATRLDLR-----RLVLAD-----GPP 965
QY 390 DNRSICALPESGCTSLPSKASTRAARLGEYLLQDLOSGOCLQ-LWDDDLGALWACPMKY 448
Db 966 H-----EPG-----PRHLDTAAG-----LIVADPRTVQPLLVRWFDDERPLPATP--- 1005
QY 449 IHKRWALVWLACLLFAAALSILLKKDHAKAAARAAAL-----LLYSADD 495
Db 1006 -----HATVATAQAALLHTRHRLDGLTEVLVDSTH 1037
QY 496 SGFERLVGALA-----SALCQLPLRVAVDLWSRRELSAOGPVAMFAQR-----QT 542
Db 1038 RRADELLAVLAEEEPSALCR-----AVERWARDERPARHRAAVTHGLRTAPHARGADRT 1092
QY 543 LOEGGVVLLFSP-----GAVALCSEWLQDGVSGPAGHGP-----DAFRA----- 583
Db 1093 LLRHAALVLLAGPSDPSPLRGALALL-----VQDPCRDRHLPAALDLFAACDPYLP 1145
QY 584 -SLSCVLPDFLQGRAPGSYVGCDFRLLHPDAVPALFR-----TVPVFTLPSQLPDLFGAL 638
Db 1146 SAVAAALPETH-----PEPVLEAFRALLGCPDAGEALRLRLADATTALT--HRVAALVGR 1198
QY 639 QOPRAPRSGRLOQRAEQVSRL--OPALDSYFHPPT-----PAPRGVGVGAGPGAGD 690
Db 1199 VTERPETAGHL---AAVYDRLDRDPAPRAVLLPLVTRLLDDGPEPARAA--LAGVLAAD 1253
QY 691 G 691
Db 1254 G 1254
```

RESULT 3

H85840

hypothetical protein yehQ [imported] - Escherichia coli (strain O157:H7, substrain ED)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: H85840

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamou, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: H85840
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-745 <STO>
A:Cross-references: GB:AB005174; NID:gl2516346; PIDN:AA057188.1; GSPDB:GN00145; UWGP:232
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yehQ

Query Match 3.1%; Score 115.5; DB 2; Length 745;
Best Local Similarity 21.4%; Pred. No. 0.83;
Matches 162; Conservative 70; Mismatches 234; Indels 291; Gaps 43;

QY 21 LERLVGPQDATH-----CSPGLSCLWSDILCLPGDIVPAGPVLAPLTHLQTELVLRC 74
DB 110 LEELATLPDATRKRAQVLVAKGIIETLF-----CAPGEIPSAIRLP-MSDVRFYRSIRF 163
QY 75 OKETDC---DCLRVAVHLAVGHWEPEDEKF-----GGAADSGVEEPRNA 119
DB 164 AR-CDCTEGTICEHV--LAVQAFVEAKTQQAETHLIWQRSEHTVSSDDPFASEGNA 220
QY 120 SLQ-----AQVLSFQA-----YPTARCVLLEVPAAALVQFGQSGSVV 159
DB 221 CROYVQQLSQALWLGISQPLIHYEAAFSRAQQAERCNR--WVSESLRQLRASV--- 274
QY 160 YDCFAA-----LGSEVRWSTPTPYEKELNHTQOL-----PALPWLNVASDG-- 203
DB 275 -DAFHARASHYHAGECLRLAALNSR-----LNCQAEMARRDSVGEVPPVWRTVVGSGIA 329
QY 204 -----DNVHLV-LNV--SEBQHEGLSLYNQVQPPKPRWKNLGPQIITLNNHTDLPV 254
DB 330 GEAKLDHLRLVSLGMRWQDIEHYGLRIWF-----TDPDGTSLIHLIS--- 371
QY 255 CLCIQVWPLEPDSVRTNICPFREDPRAHQLWQAARLLTLTQSWLL-----DAPCS 306
DB 372 -----RSPWSEQENSPAATRLFSFGAGALAGQIVSOAKRS 410
QY 307 LPAAALCWRAAGDPQCPPLV--PLSWENVTVDKVLEFPL-----LKGHPNLC 353
DB 411 ADGELLATR-----NRLSSVPLSPDAW-----RMSAPLRQPGIVALREYLRQRPSC 460
QY 354 V-----QVNSSEKLOQECL--WADSLGPLKDDVLLLETRGPQDNRSCLALEFSGCTSLP 406
DB 461 IRPLNQVDNLFILPVAECISLGWSSRQTL--DAQVISGEG-EDN--LLTL-----SLP 509
QY 407 SKASTRAA--RLGEYLLDQSGOCL-----OLWDDDD----- 436
DB 510 ASASAPYAVERMAA--LQOQDDPVLVSGFVSFVDCQGLILEPQVMTKTRAWALDAETAP 568
QY 437 -----LGALWACPMCKYIHKRWALVWLACLLFAAALSLLLLKKD--HAKAAARGAALL 490
DB 569 VVASLPSASVLPVPSTAHO-----LLMRCQALLIQLLHNGWRYQEQAIGQAEALLA 619
QY 491 YSADDSGFERLVGALASALCOLPLRVAVDLMSRRELSSAOGPVANFHAQRQTLOEGGVV 550
DB 620 NDLTAVGFYRLARVL-----GQFRNTESEARVEA 648
QY 551 LLFSPCAVALCSEWL-----ODGVSGPGAHC-----PHDAFR-----ASLSCVLPDFLO 594
DB 649 M-----NNGVLLCEBOLPPLQOQGLNRRGFGCEFCICELRDPHRRWKNKLEFLLPPE--- 702
QY 595 GRAPGSYVGACFDRLHLPDAVPALFRTVPVFTLPSQL 631
DB 703 -----EHGPAFPAL-----VDCYTSPPFL 721

RESULT 4

C99995
hypothetical protein ECs2931 [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: C99995
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C99995
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-745 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA036354.1; PID:gl3362400; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs2931

Query Match 3.1%; Score 114.5; DB 2; Length 745;
Best Local Similarity 21.6%; Pred. No. 0.99;
Matches 163; Conservative 69; Mismatches 236; Indels 287; Gaps 43;

QY 21 LERLVGPQDATH-----CSPGLSCLWSDILCLPGDIVPAGPVLAPLTHLQTELVLRC 74
DB 110 LEELATLPDATRKRAQVLVAKGIIETLF-----CAPGEIPSAIRLP-MSDVRFYRSIRF 163
QY 75 OKETDC---DCLRVAVHLAVGHWEPEDEKF-----GGAADSGVEEPRNA 119
DB 164 AR-CDCTEGTICEHV--LAVQAFVEAKTQQAETHLIWQRSEHTVSSDDPFASEGNA 220
QY 120 SLQ-----AQVLSFQA-----YPTARCVLLEVPAAALVQFGQSGSVV 159
DB 221 CROYVQQLSQALWLGISQPLIHYEAAFSRAQQAERCNR--WVSESLRQLRASV--- 274
QY 160 YDCFAA-----LGSEVRWSTPTPYEKELNHTQOL-----PALPWLNVASDG-- 203
DB 275 -DAFHARASHYHAGECLRLAALNSR-----LNCQAEMARRDSVGEVPPVWRTVVGSGIA 329
QY 204 -----DNVHLV-LNV--SEBQHEGLSLYNQVQPPKPRWKNLGPQIITLNNHTDLPV 254
DB 330 GEAKLDHLRLVSLGMRWQDIEHYGLRIWF-----TDPDGTSLIHLIS--- 371
QY 255 CLCIQVWPLEPDSVRTNICPFREDPRAHQLWQAARLLTLTQSWLL-----DAPCS 306
DB 372 -----RSPWSEQENSPAATRLFSFGAGALAGQIVSOAKRS 410
QY 307 LPAAALCWRAAGDPQCPPLV--PLSWENVTVDKVLEFPL-----LKGHPNLC 353
DB 411 ADGELLATR-----NRLSSVPLSPDAW-----RMSAPLRQPGIVALREYLRQRPSC 460
QY 354 V-----QVNSSEKLOQECL--WADSLGPLKDDVLLLETRGPQDNRSCLALEFSGCTSLP 406
DB 461 IRPLNQVDNLFILPVAECISLGWSSRQTL--DAQVISGEG-EDN--LLTL-----SLP 509
QY 407 SKASTRAA--RLGEYLLDQSGOCL-----OLWDDDD----- 436
DB 510 ASASAPYAVERMAA--LQOQDDPVLVSGFVSFVDCQGLILEPQVMTKTRAWALDAETAP 568
QY 437 -----LGALWACPMCKYIHKRWALVWLACLLFAAALSLLLLKKD--HAKAAARGAALL 490
DB 569 VVASLPSASVLPVPSTAHO-----LLMRCQALLIQLLHNGWRYQEQAIGQAEALLA 619
QY 491 YSADDSGFERLVGALASALCOLPLRVAVDLMSRRELSSAOGPVANFHAQRQTLOEGGVV 550
DB 620 NDLTAVGFYRLARVL-----GQFRNTESEARVEA 648
QY 551 LLFSPCAVALCSEWL-----ODGVSGPGAHC-----PHDAFR-----ASLSCVLPDFLO 596
DB 649 M-----NNGVLLCEBOLPPLQOQGLNRRGFGCEFCICELR--LPDHRFRWKNKLEF--- 698
QY 597 APGSYVGACFDRLHLPDAVPALFRTVPVFTLPSQL 631

R;White, O


```

Query Match      2.98; Score 108; DB 2; Length 2629;
Best Local Similarity 20.8%; Pred. No. 16;
Matches 95; Conservative 52; Mismatches 160; Indels 150; Gaps 22;

Qy 263 LEPSVRTNICPFREDPRHQNL-----WQAAERLRLTLQSWLLDAPCSLPA 309
      ||| : : : ||| :
Db 1585 LVPDLLAEHVLVYASCKPEANQKLPADVAVPHTFLRQOAKSLTLYPL--LILLOAAASQPE 1642

```

RESULT

Qy	99	PEDEEKTGGAADSGVEEPNRSALQAQQVLSFQAIPYARCVLLEVFQPAALVFGOSGVSV	158
Db	106	PEDWEAIYGEVERGETDVLIVSPERLNSVDFR-----DQVLPRLAATTGLL	151
Qy	159	VYDCFEAAALGSEVRINSYT--QPRYEKELNHQTOLPA-LPWLNVASDGDNHVLVLNVSEQ	216
Db	152	VVD---EAHCISD---WGHDFRPDYRRRLTMLAELPACVPVLATTATA-NARVTADVBSQL	205
Qy	217	HFGLSLYWNQVGQPKPRWHKHNITGPQIITPLNHTDLVPCLCIOVMPLEPDPSVRTNICBFR	276
Db	206	GTG-----AGDALVLRG-----PLDRSURLGLVLPL	232
Qy	277	EDPRAHNLMAARLLTLTQSMLLDAPCSLPAAEALCWAPRGGDPCLPLVPPLSWENVT	336
Db	233	D--AAHLRWLGERLGEL-----FCSGIITYTLTV-----AA	261
Qy	337	VDKVLEPFLKLKGHP--NLVCQVNSSBKQLQECLMADSJLGPL-----	376
Db	262	AEEITAAPLRQRGYPVASYTGKTENARDQAEDLLANRVKALVATSAIGMGFKDKPLDGFV	321

233	D	-AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	261
Qy							
237	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	265
Qy							
241	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	269
Qy							
245	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	273
Qy							
249	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	277
Qy							
253	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	281
Qy							
257	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	285
Qy							
261	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	289
Qy							
265	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	293
Qy							
269	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	297
Qy							
273	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	301
Qy							
277	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	305
Qy							
281	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	309
Qy							
285	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	313
Qy							
289	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	317
Qy							
293	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	321
Qy							
297	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	325
Qy							
301	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	329
Qy							
305	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	333
Qy							
309	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	337
Qy							
313	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	341
Qy							
317	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	345
Qy							
321	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	349
Qy							
325	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	353
Qy							
329	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	357
Qy							
333	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	361
Qy							
337	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	365
Qy							
341	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	369
Qy							
345	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	373
Qy							
349	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	377
Qy							
353	D	AAHRLAWLGERLGE-	----	PCSGIITLV	----	AA	381
Qy							
357	D	AAHRLAWLGERLGE-</					

[illegible]

A:Map position: 19ql3.4-19ql3.4

A>Note: highly expressed in colon and pancreatic cancer cells but not in the normal cell
C:Superfamily: protein-tyrosine-phosphatase, receptor type H; fibronectin type III repeat
C:Keywords: carcinogenesis; duplication; glycoprotein; phosphoprotein; phosphoric monoester
F:1-27/Domain: signal sequence #status predicted <SIG>
F:27-110/Domain: fibronectin type III repeat homology <3FNA>
F:28-1118/Product: protein-tyrosine-phosphatase, receptor type H #status predicted <MAT>
F:28-761/Domain: extracellular #status predicted <EXT>
F:116-199/Domain: fibronectin type III repeat homology <3FNB>
F:205-289/Domain: fibronectin type III repeat homology <3FNC>
F:296-379/Domain: fibronectin type III repeat homology <3FND>
F:385-468/Domain: fibronectin type III repeat homology <3FNE>
F:474-558/Domain: fibronectin type III repeat homology <3FNF>
F:564-658/Domain: fibronectin type III repeat homology <3FNG>
F:667-737/Domain: fibronectin type III repeat homology <3FNH>
F:762-778/Domain: transmembrane #status predicted <TMN>
F:779-1118/Domain: intracellular #status predicted <INT>
F:846-1070/Domain: protein-tyrosine-phosphatase homology <PTPI>
F:35,78,107,132,149,172,196,203,286,304,312,329,352,376,383,401,436,439,470,490,558,595
F:1022/Active site: Cys (phosphotyrosine intermediate) #status predicted
F:1028/Binding site: substrate phosphate (Arg) #status predicted

Query Match 2.9%; Score 107; DB 1; Length 1118;
Best Local Similarity 20.1%; Pred. No. 6.3;
Matches 88; Conservative 38; Mismatches 138; Indels 174; Gaps 21;

QY 56 ACPVPLATHLOELVLRCQKETDCDLCLRVAVHLVGHWEPEDEK-----FGGAA 109
| | | : : : | : : | : : | : : | :
DB 387 APNPV-RNLHMET-----QTNSIALC-----WEVPDGPQDYTYWVGTYG 427
| | | : : : | : : | : : | : : | :

QY 110 DSGVEEPRNASLQAQVVLVSFOAYPTARCVLLEVOVPAALVQFGOSGVSVVYDCFEAALGS 169
| | | : : : | : : | : : | : : | :
DB 428 DGGGTETRNT-----TNTSVTAERLEPGTLTYFS-----VWAENKGARGCS 467
| | | : : : | : : | : : | : : | :

QY 170 EVRIWSYTOPRYEKELNHHTQOLPALPWLNVASADGNVHLVLNVSEEQHFGLSLYNQVOG 229
| | | : : : | : : | : : | : : | :
DB 468 QNVYSISTYP-----NAVTSLSKQDWNTST-----IALRWTAPOG 502
| | | : : : | : : | : : | : : | :

QY 230 PKPKRW-----HKNLTPGOIITLNHTLDLPCLCIOWPLEPSVRTNICPFREDPRAH 282
| | | : : : | : : | : : | : : | :
DB 503 PQQSISYVSWVRGMDTPDTOSTGTD-----ITLKEAGSL-----YH 544
| | | : : : | : : | : : | : : | :

QY 283 QNLWOAARLLRLTLQSWLLDAPCSLPAEA-----LCWRAPGDPCCPLVPP 329
| | | : : : | : : | : : | : : | :
DB 545 LTVW-AERNEVRGYNETLTAA--TAPNEVTDLQNETQKNSVMWWKAPGDPHSQLYVYW 601
| | | : : : | : : | : : | : : | :

QY 330 LSWENVTVDKVLEFPLLKHGP-----NLCVQNSS-EKLQLOE-----CLWA 370
| | | : : : | : : | : : | : : | :
DB 602 VOWAS-----KGHPRRQQDPQANWNVNOTSRTNETWKVEALEPGTLNYFTVA 649
| | | : : : | : : | : : | : : | :

QY 371 DSLGPLKKDDVLLLETGRQPQDNRSILCALEPCSTSLPSKASTRAARLGVELLQDLSQGCL 430
| | | : : : | : : | : : | : : | :
DB 650 E-----RNVD-----ASTQSLCASTPYDPTVTTSCVTSAGY----- 682
| | | : : : | : : | : : | : : | :

QY 431 QLWDDLGLWACPMKY 448
| | | : : : | : : | : : | : : | :
DB 683 -----GVNLIWSCPOGGY 695
| | | : : : | : : | : : | : : | :

RESULT 12
C70841
hypothetical protein RV2102 - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: C70841
R:Cole, S.T.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
Connor, R.; Brossch, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987

A:Accession: C70841
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-238 <COL>
A:Cross-references: GB:AL021924; GB:AL123456; NID:g3261519; PIDN:CAA17285.1; PID:e1e25
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: RV2102

Query Match 2.8%; Score 104.5; DB 2; Length 238;
Best Local Similarity 25.9%; Pred. No. 1.4;
Matches 55; Conservative 21; Mismatches 93; Indels 43; Gaps 10;

Qy 495 DSGFERLVGALASALCOLPLRVAVDLWSRRRLSAGQPWAFHQAORROTLOEGGVVLLFS 554
| : | | | | | | | | | | | | | | | | | | | |
Db 29 DAG---LVTALVQSRRAPYIRIG-----IPFGKSQWAHVH--TLAENAWYAAILL 77
| : | | | | | | | | | | | | | | | | | | | |
Qy 555 PGVAALCSSEWLQDVGSPG- AHGPDAFRAISLCVLPDFLQGRAPGSYGACFDRLHPD 613
| : | | | | | | | | | | | | | | | | | | | |
Db 78 SGEM---PEDIEDVFAGLGSLFPCTARELSILDCSDPY---AVPCKHLAAATFYLLAES- 130
| : | | | | | | | | | | | | | | | | | | | |
Qy 614 AVPALFRIVPVFTLP---SQLPDFLGAQLQQPRAPSREGLEQERAEQVSALOPALDSYF-- 668
| : | | | | | | | | | | | | | | | | | | | |
Db 131 -----FDEDPFAILLAWRGREDDLNLIAARADGAAPAADHAEQVAQPLTDCLDRIYYAR 185
| : | | | | | | | | | | | | | | | | | | | |
Qy 669 -----HPPGTTPA-----PGRGVGGPGAGP 686
| : | | | | | | | | | | | | | | | | | | | |
Db 186 QADINVSPPATPTSTALLDQLPDFTGLSARGRP 217
| : | | | | | | | | | | | | | | | | | | | |

RESULT 13
S33121
homeotic protein CDP - dog (fragment)
N:Alternate names: CCAAT displacement protein; homeotic protein Clox
C:Species: Canis lupus familiaris (dog)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999
C:Accession: S33121
R:Andres, V.; Nadal-Ginard, B.; Mahdavi, V.
Development 116, 321-334, 1992
A>Title: Clox, a mammalian homeobox gene related to Drosophila cut, encodes DNA-binding
A:Reference number: S33121; MUID:93161940
A:Accession: S33121
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-975 <AND>
A:Cross-references: EMBL:X69017; NID:g2201; PIDN:CAA48782.1; PID:g2202
C:Superfamily: Homeotic protein CDP; cut repeat homology; homeobox homology
C:Keywords: DNA binding; duplication; homeobox; nucleus; transcription regulation
F:1-67/Domain: cut repeat homology (fragment) <CU1>
F:383-455/Domain: cut repeat homology <CU2>
F:566-638/Domain: cut repeat homology <CU3>
F:685-741/Domain: homeobox homology <HOX>

Query Match 2.8%; Score 104.5; DB 2; Length 975;
Best Local Similarity 20.4%; Pred. No. 8.3;
Matches 174; Conservative 74; Mismatches 244; Indels 363; Gaps 44;

Qy 3 VPWFLLSLALGRSPVVISLERLVGQDATHCS-----PGLSCRLMDSIDLCPGDIVP 55
| : | | | | | | | | | | | | | | | | | | | |
Db 190 IPSSPMSSSVSYPPPLALSLLK---PPTAPDTASTLPNPALKKEAQDALCDLPGAES 246
| : | | | | | | | | | | | | | | | | | | | |
Qy 56 AGRPVLATPHLOTLYELRCQKETDCDCLLRVAVHLAVHGHWEE-----PE 100
| : | | | | | | | | | | | | | | | | | | | |
Db 247 AQG-----VLR-----HKSELGRSGVMWKDHWSITVQPERKSAAPPE 283
| : | | | | | | | | | | | | | | | | | | | |
Qy 101 D-EKKFGGAADSG-----V 113
| : | | | | | | | | | | | | | | | | | | | |
Db 284 DAKSEEAGGTEKGCGQGCHPIASSRDPHHRRTGRNGPASPSPRTPOSSELSUTGASRS 343
| : | | | | | | | | | | | | | | | | | | | |
Qy 114 EEPRNASLQAQ--VVYLSFOAYPTARCVELLEVQVPAAALVQFGSVGVSYYDCFEAALGSEV 171
| : | | | | | | | | | | | | | | | | | | | |

A;Molecule type: DNA
A;Residues: 1-1487 <PYL>
A;Cross-references: GB:AE001274; NID:g3264850; PIDN:AAC24673.1; PID:g2266918; GSPDB:GN00
A;Experimental source: strain MHOM/IL/81/Friedlin
C;Genetics:
A;Gene: L1439.2
A;Map position: 1

Query Match 2.88; Score 104; DB 2; Length 1487;
Best Local Similarity 23.7%; Pred. No. 15;
Matches 63; Conservative 19; Mismatches 112; Indels 72; Gaps 9;
QY 444 PMDKYIHKRWALVWLACILFAALSLIILLKKDHAKAAARGRAALLLYSADDSGFERLVG 503
Db 19 PLDAQGH-----VCHARAAATAPATVVSNAEVTGARGSCAPLHASREDT--RDRDS 68
QY 504 ALASALCQLPLRVAVDLWSRR-----ELSAQGFVAMFHAQRROTLL-----OEGGVVILL 552
Db 69 PVAAPQRQOPALNAEDEWRWRASRPASSSSSSPSAODEAHSDRVVYEDGODGG----- 123
QY 553 FSPGVALCSEWLQDGVSGPGAHGPHDAFRASLSCLVLPDFLOGRAPGSYVGACFDRLHLP 612
Db 124 --EEADRSEWRHS-----GADGASDDAEASSSTA-----LHP 155
QY 613 DAVPALFRTVPVFTLPSQLPDFLGLALQOPRAPRGRGRLOERAQVSRALQPALDSYFHPFG 672
Db 156 PSPQLHQOEASPLPLSTPVTACTQQAQPP-----STTAPPPGLESDTTPPQ 203
QY 673 TPAPGRGVGP-----GAGPGAGDGT 692
Db 204 VVSOGKGTDPKRLCTAVAPASGRAT 229

Search completed: May 19, 2002, 14:11:33
Job time: 6401 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 12:24:03 ; Search time 43.31 Seconds
(without alignments)
390.268 Million cell updates/sec

Title: US-09-608-918-2
Perfect score: 3716
Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPGRGVGPGAGCGDGT 692

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA.*
1: /cgn2.6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2.6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2.6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2.6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2.6/prodata/2/iaa/PCUS_COMB.pep.*
6: /cgn2.6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	326.5	8.8	617	4	US-09-188-930-303
2	174	4.7	866	2	US-08-620-694A-10
3	174	4.7	866	3	US-09-022-255-10
4	174	4.7	866	3	US-09-022-696-10
5	174	4.7	866	3	US-08-978-773-4
6	174	4.7	866	3	US-09-022-253-10
7	174	4.7	866	3	US-09-022-260-10
8	174	4.7	866	4	US-09-022-259-10
9	174	4.7	866	4	US-09-022-257-10
10	165	4.4	864	2	US-08-620-694A-2
11	165	4.4	864	3	US-09-022-255-2
12	165	4.4	864	3	US-09-022-696-2
13	165	4.4	864	3	US-08-978-773-2
14	165	4.4	864	3	US-09-022-253-2
15	165	4.4	864	3	US-09-022-260-2
16	165	4.4	864	4	US-09-022-259-2
17	165	4.4	864	4	US-09-022-257-2
18	162	4.4	330	4	US-09-188-930-125
19	116.5	3.1	1711	2	US-08-342-930-2
20	112.5	3.0	1208	4	US-09-463-702A-2
21	108	2.9	2629	2	US-08-751-189-4
22	108	2.9	2629	2	US-09-060-836-4
23	108	2.9	2629	4	US-09-184-445-4
24	107	2.9	4928	4	US-09-036-987A-5
25	107	2.9	4928	4	US-09-370-700-5
26	106	2.9	449	3	US-08-897-236-23
27	103.5	2.8	3739	3	US-09-320-878-2

28	101.5	2.7	469	3	US-08-753-007A-8	Sequence 8, Appli
29	101.5	2.7	469	4	US-09-398-496-8	Sequence 8, Appli
30	100	2.7	954	2	US-08-749-169A-3	Sequence 3, Appli
31	100	2.7	954	2	US-09-130-032A-3	Sequence 3, Appli
32	99	2.7	3724	2	US-08-804-227C-10	Sequence 10, Appli
33	99	2.7	3724	2	US-08-804-198-4	Sequence 4, Appli
34	98	2.6	623	1	US-08-653-740-7	Sequence 7, Appli
35	98	2.6	623	2	US-09-073-594-7	Sequence 7, Appli
36	98	2.6	623	3	US-09-275-925-7	Sequence 7, Appli
37	98	2.6	2205	1	US-08-093-453B-2	Sequence 7, Appli
38	97.5	2.6	647	3	US-08-753-007A-32	Sequence 32, Appli
39	97.5	2.6	647	4	US-09-398-496-32	Sequence 32, Appli
40	97.5	2.6	3739	4	US-09-105-537-33	Sequence 33, Appli
41	97.5	2.6	11877	4	US-09-105-537-6	Sequence 6, Appli
42	97	2.6	1326	4	US-09-147-236-5	Sequence 5, Appli
43	96	2.6	405	2	US-08-483-151-4	Sequence 4, Appli
44	96	2.6	1843	4	US-09-413-814-50	Sequence 50, Appli
45	96	2.6	3729	2	US-08-804-227C-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-188-930-303
; Sequence 303, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 303
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-303

Query Match	8.8%	Score 326.5;	DB 4:	Length 617;
Best Local Similarity	26.7%	Pred. No. 2e+22;		
Matches 132;	Conservative 56;	Mismatches 184;	Indels 123;	Gaps 19;
QY	237	KNLTGQIITLNHTDLVPCLCIQVWPDPDSVRTNICPFREDPRAH-QNLWQAAR-----	290	
DB	183	KIVSGHGVDPYFELLPCMCIEASYLQEDTVRRKKCFQSWPEAYGSGFWSIRFTDYS	242	
QY	291	-----LRLLTQSLWLDAPCSLPAEALCWRAPGGDCQCPPLVPPLSWENVTYDKVLEPL	345	
DB	243	QHQNVMALTLR-----CPLKLEASLCWRDPLTPCETL-----PNATAQSEGWY	289	
QY	346	LKG-----HNNLCVQVNSSEKQLQECLEWADSLGLKDDVLLLETRGPQDNRSICALPEPSC	402	
DB	290	LENDVLHPOLCFKF-SFENSSHVEC--PHQSGSLPSTVSMDTAQO-----LTLHFSSR	341	
QY	403	TSLPSKASTRAARLG-----EYLLQDIOSSGQCLQWDDDLGAL	440	
DB	342	TYATESAAMDPLGPDTPMPVYISQFGSVPTVTLIIIPFLROENCILVWRSVDHPA	401	
QY	441	W---ACPMDKYIKRWALVWLACLLFAAALSILILIKDHAKAAARAAALLYSADDSG	497	
DB	402	WKNVLCPPD-----APYPTQLLR---SLGSGTRPVLLHHAADSEA	440	
QY	498	FERLVGALASALCQLPLRVA-----VDLWSRRELSAQGPVAFVFAORRQTLQEGGVV	549	

Db 441 ORRLVGAEL-----LFTALGGGRDVIWDLWEGTHVARIGPLPWLWARRVAREQGTV 495
QY 550 VLFSPGAVALCSEWLDQVSGPAGHGDFAFASLSCVLPDFLQGRAPSGVVGACFDRL 609
Db 496 LLLWN-----CAGPSTACSGDPQAASLRTLL-----CAAPRPLLLAYFSRL 536
QY 610 LHPDAVPALFRTVPVFLPSQLPDLFGLAQ-QPRAPRSG-----RLQERAEQVS-RA 659
Db 537 CAGKIDIPRLPRALPRYLLRDLRLLRALDAQPATLASSWSHLGAKRCLKNRLEQCHLLE 596
QY 660 LOPALDSYFHPGTP 674
Db 597 LEAAKDDYQGSTNSP 611

RESULT 2
US-08-620-694A-10
; Sequence 10, Application US/08620694A
; Patent No. 5869286
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694A
; FILING DATE: 21 MARCH 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694A
; FILING DATE: 23 MARCH 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-620-694A-10

Query Match 4.7%; Score 174; DB 2; Length 866;
Best Local Similarity 21.6%; Pred. No. 1.1e-07;
Matches 119; Conservative 76; Mismatches 189; Indels 166; Gaps 32;

QY 237 KNLIT--GPQIITLN-----HT---DLVPLCIQVWPLEPDS-----VRN-- 271
Db 66 RNLTPSPKDLQTLQHFHTQGDLPFAHIE-WTLQTDASILYLEGAELSVLQLTNTR 124

QY 272 IC---PFREDPRAHQLWQAARLRLITLQSWLIDAPCS-----LPAAALCWAPGGD 331
Db 125 LCVRFEEFLSKLRHHRRWR-----FTSFHVVDPDQYEVTVHHLPKP-----IPDGD 172
QY 322 PCOP-----LVPPLSWENVTVDKVLEPPLKG-----HPNLCOVQNSSEKLOQLQCLWADSL 373
Db 173 PNHQSKNLFVPCDCEHARKMVT-----PCMSSGSLMDPNITVETLEAHQLRVSEFTLWNES- 227
QY 374 GPLKDDVLLLETRGPQDNRSCLALEPSGCTS-----LPSKASTRAARLGEY-----LL 421
Db 228 ---THYQILLTSFPH-----MENHSCFEHMHHP-----APRPEEFHORSNTVLT 270
QY 422 QDLQS-----GQCQLQWDDDL--GALWACP-----MDKYIHKRWAL 455
Db 271 RNLKGCRRHQVQIQPFSSCL-----NDCLRHSATVSCPEMPDTPETIPDYMWLVY--WFI 325
QY 456 VWLACLLFAAALSLLILLK-----KDHAKAA-----ARGRAALLLYS 492
Db 326 TGISILLGVSVILLIVCMTWRLAGPGEKYSDDTKYTDGLPAADLIPPLPKPRKVVIIYS 385
QY 493 ADDSGFERLVGALAS-ALCOLPLRVAVDLWSRRELSAOGPVAFVFAORRROTLOEGGVVYL 551
Db 386 ADHPLYVDVVLKFAQLLTACGTEVALDLLEQAISEAGVMTWVGKQKQEMVESNKIIV 445
QY 552 LFSGCAVALCSEWLDQVSGPGA-----HGP--HDAFRASLSCVLPDFLQGRAPGSYV 602
Db 446 LCSRGTRA---KW--QALLRGAPVRLRCDHGKPVGDLFTAAMNMLPDKFKPACFGTV 500
QY 603 GACFDRLLHPDAVPALFRTVPVFLPSQLPDLFGLAQGRAPRSGRLQERAEQVSRAQLP 662
Db 501 VCYFSEVSCDGDVPLDFAAPRYPLMDREFEYVFRIQDLEMFQPGRM-HRVGELSG---- 555
QY 663 ALDSYFHPG 672
Db 556 ---DNYLRSPG 563

RESULT 3
US-09-022-255-10
; Sequence 10, Application US/09022255
; Patent No. 6072033
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,255
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/620,694
; FILING DATE: 21 MARCH 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-022-255-10

REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 866 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-255-10

Query Match 4.7%; Score 174; DB 3; Length 866;

Best Local Similarity 21.6%; Pred. No. 1.le-07;
Matches 119; Conservative 76; Mismatches 189; Indels 166; Gaps 32;

QY 237 KNL--GQIITLN---HT---DLVPCLCIQVWPLEPDS-----VRTN-- 271
Db 66 RNLTPSSPKDLQIQLHFAHTQOGDLFPVAHIE-WLQTDASILYLEGAELSVLQNTNER 124
QY 272 IC---PFREDPRAHONLWQAARLRLTLQSWLLDAPCS-----LPAEALCWRAFGD 321
Db 125 LCVREFEFLSKLRHRRWR-----FTFSHFVVDQVEYEVTVHHLKP-----IPGD 172
QY 322 PCQP-----LVPPLSWENVTVKLEPPLKLG-----HPNLCVQVNSSEKQLQECLEWADSL 373
Db 173 PNHQSKNFLVPDCEHARMKVT---PCMSGSLWDPNITVETLEAHQLRVSTFLWNES- 227
QY 374 GPLKDDVLLLETRGPDNRSLCALEPSGCTS-----LPSKASTRAARLGEY-----LL 421
Db 228 ---THYQILLTSFPH-----MENHSCFEHMHIP-----APRPEEFHQRNVTLTL 270
QY 422 ODLOS-----GQCQLQWDDDL--GALWACP-----MDKYIHKRWAL 455
Db 271 RNLKCCRHQVQIQPFSSCL---NDCLRHSAIVSCPEMPDTPPIPYMPLWY--WFI 325
QY 456 VWLACLLFAAALSILLK-----KHAKAA-----ARGRAALLYS 492
Db 326 TGISILLVGSVILLIVCMTWRLAGPGSEKYSDDTKYTDGLPAADLIPPLPKPRKVIIS 385
QY 493 ADDSGFERLVGALAS-ALCQLPLRVAVDLWSRRELSSAQGPVAFHQAOROTLQEGGVVVL 551
Db 386 ADHPLYVDVVLKFAQLTACGTEVALDLLEQAISEAGVMTWVGROKQEVESNKIIV 445
QY 552 LFSPGAVALCSEWLQDGVSGPGA-----HGP--HDAFRASLSCVLPDLQGRAPGSYV 602
Db 446 LCSRGTRA---KW--QALLGRGAPVRLRCDHGKPVGDLFTAAAMNMLLPDKRPACFGTYV 500
QY 603 GACFDRLLHPDAVPALFRTVPVFTLPSQLPDLFGALQOPRAPRSGRLOERAEQVSRALQP 662
Db 501 VCYFSEVSCDGDVPDLFGAAPRYPLMDRFEVYFRIDLEMFQPM-HRVGELSG---- 555
QY 663 ALDSYHPGP 672
Db 556 --DNYLRSPG 563

RESULT 4

US-09-022-696-10
Sequence 10, Application US/09022696
Patent No. 6072037
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6072037e1 Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle

STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,696
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 866 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-696-10

Query Match 4.7%; Score 174; DB 3; Length 866;

Best Local Similarity 21.6%; Pred. No. 1.le-07;
Matches 119; Conservative 76; Mismatches 189; Indels 166; Gaps 32;

QY 237 KNL--GQIITLN---HT---DLVPCLCIQVWPLEPDS-----VRTN-- 271
Db 66 RNLTPSSPKDLQIQLHFAHTQOGDLFPVAHIE-WLQTDASILYLEGAELSVLQNTNER 124
QY 272 IC---PFREDPRAHONLWQAARLRLTLQSWLLDAPCS-----LPAEALCWRAFGD 321
Db 125 LCVREFEFLSKLRHRRWR-----FTFSHFVVDQVEYEVTVHHLKP-----IPGD 172
QY 322 PCQP-----LVPPLSWENVTVKLEPPLKLG-----HPNLCVQVNSSEKQLQECLEWADSL 373
Db 173 PNHQSKNFLVPDCEHARMKVT---PCMSGSLWDPNITVETLEAHQLRVSTFLWNES- 227
QY 374 GPLKDDVLLLETRGPDNRSLCALEPSGCTS-----LPSKASTRAARLGEY-----LL 421
Db 228 ---THYQILLTSFPH-----MENHSCFEHMHIP-----APRPEEFHQRNVTLTL 270
QY 422 ODLOS-----GQCQLQWDDDL--GALWACP-----MDKYIHKRWAL 455
Db 271 RNLKCCRHQVQIQPFSSCL---NDCLRHSAIVSCPEMPDTPPIPYMPLWY--WFI 325
QY 456 VWLACLLFAAALSILLK-----KHAKAA-----ARGRAALLYS 492
Db 326 TGISILLVGSVILLIVCMTWRLAGPGSEKYSDDTKYTDGLPAADLIPPLPKPRKVIIS 385
QY 493 ADDSGFERLVGALAS-ALCQLPLRVAVDLWSRRELSSAQGPVAFHQAOROTLQEGGVVVL 551
Db 386 ADHPLYVDVVLKFAQLTACGTEVALDLLEQAISEAGVMTWVGROKQEVESNKIIV 445
QY 552 LFSPGAVALCSEWLQDGVSGPGA-----HGP--HDAFRASLSCVLPDLQGRAPGSYV 602
Db 446 LCSRGTRA---KW--QALLGRGAPVRLRCDHGKPVGDLFTAAAMNMLLPDKRPACFGTYV 500
QY 603 GACFDRLLHPDAVPALFRTVPVFTLPSQLPDLFGALQOPRAPRSGRLOERAEQVSRALQP 662


```
Db 501 VCYSEVSCDGDVDFLFGAAPRYPLMDRFEVYFRIQDLEMFQPGRM-HRVGELSG-----555
QY 663 ALDSYFHPGP 672
Db 556 --DNYLRSPG 563

RESULT 5
US-08-978-773-4
; Sequence 4, Application US/08978773
; Patent No. 6083906
; GENERAL INFORMATION:
; APPLICANT: Troutt, Anthony
; TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple PowerMacintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,773
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/052,525
; FILING DATE: 27 NOVEMBER 1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2623-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-978-773-4

Query Match 4.7%; Score 174; DB 3; Length 866;
Best Local Similarity 21.6%; Pred. No. 1.le-07;
Matches 119; Conservative 76; Mismatches 189; Indels 166; Gaps 32;

QY 237 KNLIT--GPQILTN-----HT--DLVPCLCIQVWPLEPDS-----VRTN-- 271
Db 66 RNLTPSPKDLQLQLHFAHTQQGDLFPVAHIE-WTQTDASILYLEGAEVSLQNTNHR 124
QY 272 IC---PFREDPRAHQLNQAARLRLTLQSWLLDAPCS-----LPAEAALCWAPGGD 321
Db 125 LCVRFELSLRHHRRW-----FTFSHFVDDPDEYEVTVHHLKP-----IPGD 172
QY 322 PCQP-----LVPPLSWENVTVDKVLEPPLKG-----HPNLCVQVNSSEKQLQECIWDLSL 373
Db 173 PNHQSKNFLVPDCEHARKVTT-----PCMSSGSLWDENITVETLEAHQLRYSFTLWNES- 227
QY 374 GPLKDDVLLLETGPGQDNLSICALEPSCGHS-----LFSKASTRAARLGEY-----LL 421
Db 228 ----THYQILLTFFPH-----MENHSCFEHMHIP-----APRPEEFHQRNSVTLTL 270
QY 422 QDLQS-----GQCILQLWDDDL--GALWACP-----MDKYIHKRWAL 455
```

```
Db 271 RNLKGCRRHQVQIQPFSSCL-----NDCLRHSAVSCPEMPDTPPEIPDPYMLWY--WFI 325
QY 456 VWLACLLFAAALSILLK-----KDHAKEA-----ARGRAALLYS 492
Db 326 TGISILLVGSVILLIVCMTRLAGPGSEKYSDDTKYTDGLPAADLLPPPLKPKRWIYS 385
QY 493 ADDSGFERLVGALAS--ALCOLPLRVAVDLWSRRELQAQGPVAVFHAQRROTLOEGGVVVL 551
Db 386 ADHPLYVDVVLKFAQFLLTACGTEVALDLLEQAISEAGVMTWVGKQKQEMVESNKIIV 445
QY 552 LFSPPGAVALCSEWLQGVSPGA-----HGP--HDAFRASLSCVLPDFLOQRAFGSYV 602
Db 446 LCSRGTRA---KW--QALLRGAPVRLRCDHGKPGVGLFTAAANNMILPDFKRPACFTYV 500
QY 603 GACFDRLLHPDAVPALFRTVPVFTLPSQLPDFLGAQQPRAPSGRLQERAEVSRALQP 662
Db 501 VCYFSEVSCDGDVDFLFGAAPRYPLMDRFEVYFRIQDLEMFQPGRM-HRVGELSG-----555
QY 663 ALDSYFHPGP 672
Db 556 --DNYLRSPG 563

RESULT 6
US-09-022-253-10
; Sequence 10, Application US/09022253
; Patent No. 6096305
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,253
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694
; FILING DATE: 21-MARCH-1996
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/410,535
; FILING DATE: 23 MARCH 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

US-09-022-253-10

Query Match 4.7%; Score 174; DB 3; Length 866;
Best Local Similarity 21.6%; Pred. No. 1.le-07;
Matches 119; Conservative 76; Mismatches 189; Indels 166; Gaps 32;

QY 237 KNLTL-GPOITLNL---HT---DLVPCICIQVWPLEPDS-----VRTN-- 271
DB 66 RLTPSSPKDQIQIHFHTAQOGDLFPVAHIE-WTQTDASILYLEGAELSVLQNTN 124
QY 272 IC---PFREDPRAHONLWQAARLLTLQSWLLDAPCS-----LPAAALCWAPG 321
DB 125 LCVREFELSKLRHHRRWR-----FTFSHFVVDPDQEVETVHHLKPK-----IPDGD 172
QY 322 PCQP-----LVPPLSWENTYVKVLEFFLLKG-----HPNLCVQVNSSEKQLQCLNADSL 373
DB 173 PNHQSKNFLVPDCEHARKMKT---PCMSGSLMDPNITVETLEAHQLRVSTLWNES- 227
QY 374 GPLKDDVLLLETRGPQDNRSICALPESGCTS-----LPSKASTRAARLGEY-----LL 421
DB 228 ---THYQILLTSPH-----MENHSCFEHMHIP-----APRPEEFHQRSNVTLL 270
QY 422 ODLOS-----GQCLQWLDDDL--GALWACP-----MDKYIHKRWAL 455
DB 271 RNKAGCCRHQVQIOQFFSSCL---NDCLRHSAVSCPEMPDTPPEIPDYMLWVY--WFI 325
QY 456 VWLACLLFAAALSLLILK-----KDHAKAA-----ARGRAALLYS 492
DB 326 TGISILLVGSVILLIVCMTWRLAGPGESEKYSDDTKYTDGLPAADLIPPPKPKVWIIYS 385
QY 493 ADDSGFERLVGALAS-ALCQLPLRVAVDLWSRRELSAQPVAWFHAORRQTLQEGGVVL 551
DB 386 ADHPLYVDVLKFAQFLLTACGTEVALDLLEEQAISEAGVMTWVGKQEMVESNKKIIV 445
QY 552 LFSFGAVALCSEWLQDGVSGPGA-----HGP---HDAFRASLSVLPDFLQGRAGSYV 602
DB 446 LCSRTRA---KW---QALLGRGAPVRLCDHGKPVGDLFTAANNMILPDFKRPACFTYV 500
QY 603 GACFDRLLHPDAVPALFRTVFTLPSQLPDFLQALQOPRAPSRLQERAEQVSRALQP 662
DB 501 VCFSEVSCDGVDPDLFGAAPRYPLMDREFEYVRIQDLEMFQGRM-HRVGELSG-----555
QY 663 ALDSYFHPGP 672
DB 556 ---DNYLRSPG 563

RESULT 7

US-09-022-260-10
Sequence 10, Application US/09022260
Patent No. 6100235

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin
APPLICANT: Springs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,260
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 866 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-260-10

Query Match 4.7%; Score 174; DB 3; Length 866;

Best Local Similarity 21.6%; Pred. No. 1.le-07;

Matches 119; Conservative 76; Mismatches 189; Indels 166; Gaps 32;

QY 237 KNLTL-GPOITLNL---HT---DLVPCICIQVWPLEPDS-----VRTN-- 271
DB 66 RLTPSSPKDQIQIHFHTAQOGDLFPVAHIE-WTQTDASILYLEGAELSVLQNTN 124
QY 272 IC---PFREDPRAHONLWQAARLLTLQSWLLDAPCS-----LPAAALCWAPG 321
DB 125 LCVREFELSKLRHHRRWR-----FTFSHFVVDPDQEVETVHHLKPK-----IPDGD 172
QY 322 PCQP-----LVPPLSWENTYVKVLEFFLLKG-----HPNLCVQVNSSEKQLQCLNADSL 373
DB 173 PNHQSKNFLVPDCEHARKMKT---PCMSGSLMDPNITVETLEAHQLRVSTLWNES- 227
QY 374 GPLKDDVLLLETRGPQDNRSICALPESGCTS-----LPSKASTRAARLGEY-----LL 421
DB 228 ---THYQILLTSPH-----MENHSCFEHMHIP-----APRPEEFHQRSNVTLL 270
QY 422 ODLOS-----GQCLQWLDDDL--GALWACP-----MDKYIHKRWAL 455
DB 271 RNKAGCCRHQVQIOQFFSSCL---NDCLRHSAVSCPEMPDTPPEIPDYMLWVY--WFI 325
QY 456 VWLACLLFAAALSLLILK-----KDHAKAA-----ARGRAALLYS 492
DB 326 TGISILLVGSVILLIVCMTWRLAGPGESEKYSDDTKYTDGLPAADLIPPPKPKVWIIYS 385
QY 493 ADDSGFERLVGALAS-ALCQLPLRVAVDLWSRRELSAQPVAWFHAORRQTLQEGGVVL 551
DB 386 ADHPLYVDVLKFAQFLLTACGTEVALDLLEEQAISEAGVMTWVGKQEMVESNKKIIV 445
QY 552 LFSFGAVALCSEWLQDGVSGPGA-----HGP---HDAFRASLSVLPDFLQGRAGSYV 602
DB 446 LCSRTRA---KW---QALLGRGAPVRLCDHGKPVGDLFTAANNMILPDFKRPACFTYV 500
QY 603 GACFDRLLHPDAVPALFRTVFTLPSQLPDFLQALQOPRAPSRLQERAEQVSRALQP 662
DB 501 VCFSEVSCDGVDPDLFGAAPRYPLMDREFEYVRIQDLEMFQGRM-HRVGELSG-----555
QY 663 ALDSYFHPGP 672
DB 556 ---DNYLRSPG 563

RESULT 8

US-09-022-259-10
Sequence 10, Application US/09022259

APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,255
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/620,694
FILING DATE: 21 MARCH 1996
APPLICATION NUMBER: USSN 08/538,765
FILING DATE: 7 AUGUST 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-255-2

Query Match 4.4%; Score 165; DB 3; Length 864;
Best Local Similarity 19.8%; Pred. No. 8.2e-07;
Matches 123; Conservative 79; Mismatches 209; Indels 210; Gaps 31;
QY 192 PALPW---LNVSDGDNVHLVNL-----VSEQHFGLSLYWNQVQGP--KPRW--HKNL 239
Db 14 PALGWLILLNLVLAFGASPRLLDFPAPVCAQE--GLSC---RVKNSTCLDDSWIHPKNL 68
QY 240 T--GPQITFLN-----HTDLVPCLCIQVWPLEPDS-----VRTN--IC- 273
Db 69 TPSSPNTIYINLSVSTQGEVLVPLVHVE-WLQTDASILYLEGAELSVLQNTNERLCV 127
QY 274 --PFREDPRAHNLQAARLRLTLQSLWLDAPCSLPAAE-----ALCWAPGDDP--- 322
Db 128 KFQFLSMLQHHRKWR-----FSESHFVVD-----PGQYEYVTVHHLPKPIPDGDNHK 176
QY 323 -----CQPLVPLSWENVTVDKLEPPLKLGHNLCVQVNSSEKLQOECLWADSLG 374
Db 177 SKIIFVPCDESKMKTTSVSSGSLWD-----PNITVETLDTQHLRVDFTLWNEST- 228
QY 375 PLKDDVLLLETRGPQDNRSICALPESGCTSL-----PSKASTRAARLGEYLLQDLSQG 428
Db 229 PIQ---VLESFSDSENHS-----CFDVVKQIFAPRQEEFHQRANVTFTLSKFH--- 274
QY 429 CLQLWDDDLGALWAC-----PMDKYIHKR 452
Db 275 -----WCCHHHVQVQPFSSCLNCLRHAVTVPVCPVISNTVTPKPVADTI-PL 321
QY 453 WA---LVNWLACLLFAAALSLILLK-----RDHAKAAARGRA 486

Db 322 WYUGITLTIALLVGSVIVLIICMTWRLSGADQEKHGDDSKINGILPVADLTPTPLRPRK 381
QY 487 ALLYSADDSGFERLVGALASAL-CQLPLURVAVDLWSRRELSAQGPVAFHQAQRROTLOE 545
Db 382 WVIVISADHPLVEVVLKFAQLITACGTEVALDLLEEQVISEVGVMTWWSKQKQDMVES 441
QY 546 GGVVYLLFSPG-----AVAL-CSEWLQDGVSGPCAHGPHDAFRASLSCVLP 590
Db 442 NSKIILCSRGTOAKWKAILGWAEPVQLRCDHWKPAQ-----DLFTAAMNMLP 491
QY 591 DFLQGRAPGSYVACGDFRLLHDPALPRTVPTVFTLPSQLPDLFGALQOQPRAPRGRLO 650
Db 492 DFKRPACFGTYVVCYFSGICSERDVPDLFNITSRYPLMDRFEVYFRIQDLEMFEPGRMH 551
QY 651 ERAEQVSRALQPALDSYFHP 671
Db 552 HVRELTG-----DNYLOSP 565

RESULT 12
US-09-022-696-2
Sequence 2, Application US/09022696
Patent No. 6072037
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,696
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-696-2

Query Match 4.4%; Score 165; DB 3; Length 864;
Best Local Similarity 19.8%; Pred. No. 8.2e-07;
Matches 123; Conservative 79; Mismatches 209; Indels 210; Gaps 31;

QY	192	PALPW----	LNWSADGDNVHLVLN----	VSREQHFGLSLYWNOVQGP--	KPRW--	HKML	239	
Db	14	PALGWLLLNVLN	PAGRASPLRDLFPAPCAQE--	GLSC----	RVKNSTCLDSDWIHPKRL	68		
QY	240	T--GPOIITLN----	HTDLVPCICIQWPLEPDS----	VRTN--	IC--	273		
Db	69	TPSSPKNIYINLSV	SQTQHGELPVLUHVE--	WTJQTDASILYLEGAELS	VLQJNTNERUCV	127		
QY	274	--PFREDPRAHQNL	WQAARLLTLTQSLWLDAPCSLPAEA----	ALCWRA	PAGDGP----	322		
Db	128	KFOFLMQLHRRKRW----	FSPSHFVVD-----	PGQEYEVTVHHLPKPI	PDGDPNHK	176		
QY	323	-----CQPLVPPLSWEN	VTVDKVLEPPLKGHPNLQVQNSSEK	LQLOECLWADSLG	374			
Db	177	SKIIFVPDCEDSK	NKMTTSCVSSGLWD-----	PNITVETLDTQHLRV	DFTLWNEST--	228		
QY	375	PLKDDVLLITRGP	QPNRSLCALEPSGCTSL-----	PSKASTRAARLGEYL	LLODLOSOQ	428		
Db	229	PYO-----VLLSE	FSDSNHS-----	CFDVVKQIFAPHOEE	FORANVTFTLSKFH----	274		
QY	429	CLQLWDDDLGAL	MAC-----	WCCHHHVQVQPFSS	CLNDCLRHA	VTVP	452	
Db	275	-----WCCHHHVQVQ	PFFSSCLNDCLRHA	VTVP	CVISNTVTVPK	PVADYI--	321	
QY	453	WA---LVM	LACLFAPAAALSILILLK-----	KDHAKA	AARGRA	486		
Db	322	WVYGLTILAI	LLVGSIVILIICTMWR	LSGADQEKHGDDSK	INGILPVADLT	PPPLRPRK	381	
QY	487	ALLYSADDSG	FRLVGALASAL--	COLPLRVAVD	WLSRRELS	NOGVPANFHA	QROTLQE	545
Db	382	WVIYSADHP	PLYVEVWLKFAQLIT	ACGTEVALDLLEQ	VIIEGVVMT	VWVRKQEMVES	441	
QY	546	GGVVWLLFSPG-----	-----	AVAL--	CSEWLQDVGSG	PGAHPHDAFRAS	LSCVLP	590
Db	442	NSKIIILCS	RGTKAKWAILGWA	EPVQVLRCD	HWKPAG-----	DLFTAAMN	MIPL	491
QY	591	DFLOGRAPG	SYVGACFDRLLHPDAV	PALFRFTVFTLP	SQPLD	GALQOPRAPS	RGRLQ	650
Db	492	DFKRPAC	FGYVVVYFSGIC	SERDVPDLFNITS	RYPLMDR	FEVEVYFRI	QOLEMFEPGRMH	551
QY	651	ERABQVS	RALQPLADSYFHP	671				
Db	552	HVRELTG-----	-----	DNYLOSP	565			

RESULT 13
US-08-978-773-2
Sequence 2, Application US/08978773
Patent No. 6083906
GENERAL INFORMATION:
APPLICANT: Trout, Anthony
TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple PowerMacintosh
OPERATING SYSTEM: Apple Operating System 7.5.1
SOFTWARE: Microsoft Word for PowerMacintosh, Version. 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,773
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/052,525
FILING DATE: 27 NOVEMBER 1996

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2623-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 864 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-978-773-2

Query Match 4.4%; Score 165; DB 3; Length 864;
Best Local Similarity 19.8%; Pred. NO. 8.2e-07;
Matches 123; Conservative 79; Mismatches 209; Indels 210; Gaps

RESULT 14
US-09-022-253-2
; Sequence 2, Application US/09022253
; Patent No. 6096305
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie

```

Qy 453 WA---LWWLACLLFAAALSLLILLK-----KDHAKAAAARGRA 486
      |      :|      :|      :|      :|      :|      :|
Db 322 WYGLITLILIAILLGVSVTLIICMTWRLSGAQDEKHGDSKINGILPVAULTPPPLRPKR 381
      |      :|      :|      :|      :|      :|      :|
Qy 487 ALLIYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSAQGPVAFHAORRQLQE 545
      :|      :|      :|      :|      :|      :|      :|
Db 382 VWIYSADHPDIYEVVLKFAQFLITACGTEVALDLLEEQVISEGVNMTWVSROKQEWVES 441
      :|      :|      :|      :|      :|      :|      :|
Qy 546 GGVVLLFLSPG-----AVAL-CSEWLQDGVSGPGAHPDHFASLSLSCVLP 590
      :|      :|      :|      :|      :|      :|      :|
Db 442 NSKIILCSRGTOAKKAILGWAEPVQLRCDHKKPAG-----DLFTAAMNMLP 491
      :|      :|      :|      :|      :|      :|      :|
Qy 591 DFLOGRAGPSVVGACFBRLLHPDPAVPALFTVPVFTLPSQLPDELGAQQPAPRPSGRLO 650
      :|      :|      :|      :|      :|      :|      :|
Db 492 DFKRPACFGTYVWCYFGSICSERVDPDLFNITSRYPLMDRFEVYFRIQDLEMFEPGRMH 551
      :|      :|      :|      :|      :|      :|      :|
Qy 651 ERAEQVSRALQPALDSYFHP 671
      |      :|      :|
Db 552 HVRELTG-----DNYLQSP 565
      :|      :|      :|

RESULT 15
US-09-022-260-2
; Sequence 2, Application US/09022260
; Patent No. 6100235
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1

```

APPLICATION NUMBER: US/09/022,360
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/620,694
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 08/410,535
 FILING DATE: 23 MARCH 1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,695
 REFERENCE/DOCKET NUMBER: 2617-B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206)587-0430
 TELEFAX: (206)
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 864 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-022-260-2

Query Match 4.4%; Score 165; DB 3; Length 864;
 Best Local Similarity 19.8%; Pred. No. 8.2e-07;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 13:26:44 ; Search time 115.29 seconds
(without alignments)
1038.360 Million cell updates/sec

Title: US-09-608-918-2
Perfect score: 3716
Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPRGVGPAGAGDGT 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2598	69.9	538	4 Q9BR97	Q9br97 homo sapien
2	1812.5	48.8	567	11 Q99J43	Q99j43 mus musculu
3	174	4.7	866	4 O43844	O43844 homo sapien
4	173	4.7	866	4 Q96F46	Q96f46 homo sapien
5	165	4.4	864	11 Q60943	Q60943 mus musculu
6	131.5	3.5	4340	2 Q30764	Q30764 streptomyc
7	130.5	3.5	1366	2 Q92529	Q92529 streptomyc
8	116	3.1	1021	4 Q96M86	Q96m86 homo sapien
9	114.5	3.1	745	2 Q9EYF6	Q9eyf6 escherichia
10	114.5	3.1	1193	16 Q9HZE5	Q9hze5 pseudomonas
11	112.5	3.0	657	4 Q96DW2	Q96dw2 homo sapien
12	112.5	3.0	744	4 Q96F55	Q96f55 homo sapien
13	111.5	3.0	478	16 Q9RY89	Q9ry89 deinococcus
14	111	3.0	1279	12 Q66031	Q66031 cercopithec
15	110	3.0	757	6 Q9N015	Q9n015 macaca fasc
16	109.5	2.9	620	4 Q9HCN6	Q9hcn6 homo sapien

17	109	2.9	427	2 Q9SIL6	Q9sll6 streptovirt
18	109	2.9	603	12 P89140	P89140 pseudorabie
19	109	2.9	604	12 Q9YVB7	Q9yvb7 pseudorabie
20	109	2.9	1343	12 Q05635	Q05635 bovine herp
21	108.5	2.9	1072	4 Q96EP0	Q96ep0 homo sapien
22	108	2.9	361	2 O86440	O86440 pseudomonas
23	108	2.9	2629	11 P97499	P97499 mus musculu
24	108	2.9	10917	2 Q93NW6	Q93nw6 streptomyc
25	107	2.9	719	2 O69992	O69992 streptomyc
26	107	2.9	825	4 Q96P01	Q96p01 homo sapien
27	107	2.9	1001	2 Q9KZM8	Q9kzm8 streptomyc
28	107	2.9	1118	4 Q15426	Q15426 homo sapien
29	107	2.9	2003	4 O00306	O00306 homo sapien
30	107	2.9	4928	2 Q9ALM3	Q9alm3 saccharopol
31	106.5	2.9	679	2 Q9F2P6	Q9f2p6 streptomyc
32	106	2.9	1211	11 Q9R0L3	Q9r0l3 rattus norv
33	106	2.9	1235	11 Q9R0L5	Q9r0l5 rattus norv
34	106	2.9	1273	11 Q9R0L4	Q9r0l4 rattus norv
35	106	2.9	1531	5 Q9W0P8	Q9w0p8 drosophila
36	106	2.9	9510	2 Q93NX9	Q93nx9 streptomyc
37	105.5	2.8	1354	11 Q9EPW8	Q9epw8 mus musculu
38	105.5	2.8	1542	4 O15035	O15035 homo sapien
39	105	2.8	1766	12 Q9J599	Q9j599 fowlpox vir
40	104.5	2.8	238	16 O53500	O53500 mycobacteri
41	104	2.8	1256	11 Q99M76	Q99m76 rattus norv
42	104	2.8	1487	5 O15843	O15843 leishmania
43	103.5	2.8	2055	4 O75055	O75055 homo sapien
44	103	2.8	414	4 Q9H612	Q9h612 homo sapien
45	103	2.8	643	12 Q85050	Q85050 pseudorabie

ALIGNMENTS

RESULT 1

Q9BR97

ID Q9BR97 PRELIMINARY; PRT; 538 AA.

AC Q9BR97; (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 59.1 KDA PROTEIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-ENDOMETRIAL ADENOCARCINOMA;

RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC006411; AAH06411.1; -

KW Hypothetical protein.

SQ SEQUENCE 538 AA; 59127 MW; 65E6344DA6A5AFD2 CRC64;

Query Match 69.9%; Score 2598; DB 4; Length 538;
Best Local Similarity 99.4%; Pred. No. 8e-208;
Matches 480; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MPVPWFLLSLALGRSPVLSRLVGVGDATHCSFGLSCRLWSDILCLPGDIVPARGPV 60

Db 1 MPVPWFLLSLALGRSPVLSRLVGVGDATHCSFGLSCRLWSDILCLPGDIVPARGPV 60

Qy 61 LAPHLTQELVLRCKETDCDCLRVAVHGVHWEPEDEEFKGAADSGVEPRNAS 120

Db 61 LAPHLTQELVLRCKETDCDCLRVAVHGVHWEPEDEEFKGAADSGVEPRNAS 120

Qy 121 LQAQVLSFQAYPTARCVCVLEVVQVPAALVQFGSGSVVYDCFEAALGSEVRIMSYQPR 180

Db 121 LQAQVLSFQAYPTARCVCVLEVVQVPAALVQFGSGSVVYDCFEAALGSEVRIMSYQPR 180

Qy 181 YEKELNHTOOLPALPWLNLVNSADGDNVHLVNLVNSEQHFGLSLYWNQVQPPKPRWKNLT 240

```

Db 181 YEKLNHTQOLPALPNLWNSADGDNVHLVNLVSEQHFGLSLYWNQVGGPKPRWHKNT 240
Qy 241 GPQIITLNTDLVPCLCIQWPLEPDSVYRTNCPFPREDPRAHONLWQAAARLLTLTQSWL 300
Db 241 GPQIITLNTDLVPCLCIQWPLEPDSVYRTNCPFPREDPRAHONLWQAAARLLTLTQSWL 300
Qy 301 LDAPCSLPAEALCWAPGDCQPLVPLSWENVTVDKYLEPFLKKGHPNLCVQVNSSE 360
Db 301 LDAPCSLPAEALCWAPGDCQPLVPLSWENVTVDKYLEPFLKKGHPNLCVQVNSSE 360
Qy 361 KLQQLQECWLADSLGPKLDDVLLLETRGPQDNRSICALPESGCTSLPSKASTRAARLGEYL 420
Db 361 KLQQLQECWLADSLGPKLDDVLLLETRGPQDNRSICALPESGCTSLPSKASTRAARLGEYL 420
Qy 421 LDQDLSGQCLQLWDDDLGALWACPMCKYTHKRWALVWLACLLFAAALLSLTLKKDHAKA 480
Db 421 LDQDLSGQCLQLWDDDLGALWACPMCKYTHKRWALVWLACLLFAAALLSLTLKKDHAKG 480
Qy 481 AAR 483
Db 481 WLR 483

RESULT 2
Qy 999J43 ID Q99J43 PRELIMINARY; PRT; 567 AA.
AC Q99J43 DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 62.8 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004759; AAH04759.1;
KW Hypothetical protein.
SQ SEQUENCE 567 AA; 62798 MW; C1AAAB79E2006B1D CRC64;

```

```

Query Match 48.8%; Score 1812.5; DB 11; Length 567;
Best Local Similarity 62.9%; Pred. No. 21e-142;
Matches 357; Conservative 50; Mismatches 130; Indels 31; Gaps 7;

Qy 1 MPVPFLLSLALGRSPVVLSERLVGPQDATHCSGLSCLRLWDSILCLPGDIVPAGPV 60
Db 1 MPVSNFLSLALGRNPVVVSLERLMEPQDARCSGLSCHLWDGDLCLPGSLQSAAGPV 60

Qy 61 LAPTHQLTELVRCKQETDCDLCRLVAVHLVHGWEEPEDEKFGGAADSGVEEPRNAS 120
Db 61 LVPTRLQTELVRCKQETDCDLCRLVAVHLVHGWEEPEDEKFGGAADSGVEEPRNAS 116

Qy 121 LQAQVVLFOAYPTARCVLLEVOVPAALVQFGSGSVYDCEFAALGSEVRITWSTQPR 180
Db 117 LQAQVVLFOAYPIARCALLEVOVPAALVQFGSGSVYDCEFAALGSEVRITWSTQPR 176

Qy 181 YEKLNHTQOLP-----ALPWLNVSDGDNVHLVNLVSEQHFGLSLYWN 225
Db 177 YQKELNLTQOLPDCRGLEVRDSIQSGWLPWLNVSTDGDGNVLTLDLSEQDFSLYLRL 236

Qy 226 QVQGPKPRWHKNTGPOIITLNTDLVPCLCIQWPLEPDSVYRTNCPFPREDPRAHONL 285
Db 237 PVPDALKSLWYKNTGPNQITLNTDLVPCLCIQWPLEPDSVYRTNCPFPREDPRAHONL 296

Qy 286 WQAAARLLTLTQSWLDDPCSLPAEALCWAPGDCQPLVPLSWENVTVDKYLEPFL 345
Db 297 WHIARLRLVSPGVMLDAPCCLPGLKVTLCWQAPDQSPCQPLVPPVPQKNATVNEPDQFL 356

```

```

Qy 346 LKGHNPNCVQVNSSEKQLQECWLADSLGPKLDDVLLLETRGPQDNRSICALPESGCTSL 405
Db 357 VAGHPNLCVQVSTWEKVQLQACWLADSLGPKLDDVLLLETRGPQDNRSICALPESGCTSL 416
Qy 406 PSKASTRAARLGEYLQDLSGQCLQLW- DDDLGALWACPMCKYTHKRWALVWLACLLFA 464
Db 417 PSMASTRAARLGEELQDLSGQCLQLW- DDDLGALWACPMCKYTHKRWALVWLACLLFA 476
Qy 465 AALSILILLKKDHAKAAAGRAALLLYS-----ADDSGFERLVGALASALCOLPLRVAV 518
Db 477 AALSILILLKKDHAKAAAGRAALLLYS-----ADDSGFERLVGALASALCOLPLRVAV 518
Qy 519 DLWSRRELQAQGVPAWFAHQRRTLQEG 546
Db 534 PRSSPCPRSCR--LSWMHCREAAPLPRG 559

RESULT 3
Qy 043844 ID 043844 PRELIMINARY; PRT; 866 AA.
AC 043844 DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE IL-17 RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98035683; PubMed=9367539;
RA Yao Z., Spriggs M.K., Derry J.M.J., Strockbine L., Park L.S.,
VandenBos T., Zappone J., Painter S.L., Armitage R.J.;
RT "Molecular characterization of the human interleukin (IL)-17
receptor.";
RL Cytokine 9:794-800(1997).
DR EMBL; U58917; AAB99730.1;
KW Receptor.
SQ SEQUENCE 866 AA; 96122 MW; 88AF626A83F3FF70 CRC64;

```

```

Query Match 4.7%; Score 174; DB 4; Length 866;
Best Local Similarity 21.6%; Pred. No. 8.6e-06;
Matches 119; Conservative 76; Mismatches 189; Indels 166; Gaps 32;

Qy 237 KNLT--GPQIITLN-----HT---DLVPCLCIQWPLEPDS-----VRTN-- 271
Db 66 RNLTPSSPKDLQIQLHFAHTQOGLFPVAHTE-WTLQTDASILYLEGAELSVLQLTNTER 124

Qy 272 IC---PREDPRAHONLWQAAARLLTLTQSWLDDAPCS-----LPAAALCWAPGCD 321
Db 125 LCVRFELSLKLRHHRWR-----FTFSHEVVDQDQYEVTVHHLPRK-----IPGD 172

Qy 322 PCQP---LVPLSWENVTVDKLEFFLLKG---HPLNLCVQVNSSEKQLQECWLADSL 373
Db 173 PNHOSKNFLVPCDCHARKVTT---PCMSSGSLWDPNITVETLEAHLQVSLFVWNES- 227

Qy 374 GPLKDDVLLLETRGPQDNRSICALPESGCTSL---LPSKASTRAARLGEY-----LL 421
Db 228 ---THYQILLTSFPH-----MENHSCFEMHHP-----APRPEFHQRSNVTUTL 270

Qy 422 QDLOS-----GQCLQLWDDDL--CALWACP-----MDKYTHKRWAL 455
Db 271 RNLKGCGRHQVQIQPFSSCL---NDCLRHSATVSCPEMDPTPEIPDYMPLWY--WFI 325

Qy 456 VWLACLFLFAALSLILLK-----KDHAKA-----ARGAALLLYS 492
Db 326 TGISILLVGVILLIVCMTWRLAGPSEKYSDDTKYTDGLPAADLIPLPKRKKWIIYS 385

Qy 493 ADSGFERLVGALAS--ALCOLPLRVAVDLMSRRRELSAQGPVAFHQRRTLQEGVVVL 551
Db 386 ADHPLYVDVVLKFAQLLTACGTEVALDLLEEQAISEAGVMTWVGRQKQWVESNKIIV 445

```

OY	552	LFSGVAALCGSEWJLDDGVSGGA-----HGP--HDAFRASISCVLPDLQRAQSSVY	602
Dd	446	LCSSKSTRA--KW--QALLGGAGAVRRLCHDGKRGVGLFTFAAMMILPDKRPACTETV	500
OY	603	GACFDRLHPDVALEFRTVEVTFLPSQLDPFLIGALQOPRAPRSGRLOERAQOVSRALQP	662Z
Dd	501	VCEYSFEVSCDDGVLDPLFGAARVPYLDMRFEEVYPRIODLEMFGQRW-HRVGELSG----	555S
OY	663	ALDSVFHPRG	672
Dd	556	--DNLTLSPG	563
 RESULT 4			
ID	O96F46	PRELIMINARY;	PRT; 866 AA.
AC	O96F46;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	SIMILAR TO INTERLEUKIN 17 RECEPTOR.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=UTERUS, AND LEIOMYOSARCOMA;		
RA	Strausberg R.;		
RL	Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL, BC011624, AHH11624.1; -.		
KW	Receptor.		
SO	SEQUENCE	866 AA;	96131 MW; 28330BED2303BDC9 CRC64;

```

Query Match Similarity      4.7%; Score 173; DB 4; Length 866;
Best Local Similarity      21.3%; Pred. No. 1e-05;
Matches 117; Conservative  76; PmedMatches 191; Indels 166; Gaps 31.

QY      237 KNLTT-GPQIITLN-----HT---DLVPCICIQVMPLEPPS-----VFTN-- 271
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db      66 RNLFTSSPRDQIQHNFATITQGGDLEPRVANHIE-WLQDTASTILYLEGALSVLQDNTNR 124
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY      272 IC---PFREDPRANOIWMQARLRLLITQSWILDAPCS-----LPRAALCWRAPGD 321
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db      125 LCVRFEFSLKLNHNHRMR-----FTSFHFVDPDQEXEYVTHNLRK-----IPDGD 172
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY      322 PCQF---LTPRLSMENYTVDKVLERPILKG-----HPIICQVQNSSEKIQIOESCLMADBL 373
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db      173 PNHQSKNFELPQCEHARMKVTT-----PCMSGSLMDPNTVTELEHNOIRVSFTLNWES- 227
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY      374 GPLKDVLLLETRCFQPDNRSICALERPSCGTS---LPKASGRARALRGY-----TL 421
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db      228 -----THYQILITSPH-----MEHNSCFEIHNIHP-----APRPEEFHQRSNTLT 270
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY      422 ODLOS-----GQCILMDDDL--GALMACP-----MDKYIHKRMAL 455
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db      271 RNLKCCCRHQVQIOPEFFSSCL---NDCLRHSAIVSCPEMRDPREPIRYDNPMLVY--WPI 325
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY      456 VWLACLLFPAALSLTILLK-----KDKAKKAARGRAALLYS 492
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db      326 TGISILLVSVLITLYCMTWRMLRAGSESEKYSDTKYTDGLPYVADLIRPLRKPRKWIYS 385
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY      493 ADDSGFERLYGALAS-ALCQLPLRAYAVDMSRELISAQGVPMFHNQROTLQEGGVVL 551
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db      386 ADHPLYVDVYLFAOFLIRACGT EYALDLLEDAIS EAGVMTWVGKQKQOMESNSKITV 445
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY      552 LFPSPAVALCSMLQDGVSGPGA-----HGP--HDAERASTLCVLPFLQGRAGSGTV 602
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db      446 LCSRGTGRA---RW--QALLRGARPYRLCRDCHKRPVQDLETTAAMNMNLTLPDFKRRACGTYV 500
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

QY      603 GACFDRLLHPDAVPALEFRTVPVFTLPSQLPDFLGALQDRPARPSGRLOERAEVNSRALQP 662
      :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||

```

Db	501	VCYSEVSCDGDVDFLFAAPRYFLMDRFEEVYERIDDLFMFGGR-NHRVGLSG----	555
Oy	663	ALDSTFHPG 672	
Db	556	--DNVLRSPG 563	
RESULT	5		
Q60943			
ID	Q60943	PRELIMINARY;	PRT; 864 AA.
AC	Q60943;		
DT	01-NOV-1996 (TReMBLrel. 01, Created)		
DT	01-NOV-1996 (TReMBLrel. 01, Last sequence update)		
DT	01-DEC-2001 (TReMBLrel. 19, Last annotation update)		
DE	INTERLEUKIN 17 RECEPTOR.		
GN	IL17R.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=THYMOMA EL4;		
RX	MEDLINE=96111968; PubMed=8777726;		
RA	Yao Z., Fanslow W.C., Seldin M.F., Rousseau A.M., Painter S.L.,		
RT	ComEAU M.R., Cohen J.I., Spriggs M.K.;		
RT	"Hepesvirus Salmelit encodes a new cytokine, IL-17, which binds to a		
RT	novel cytokine receptor."		
RL	Immunity 3:811-821(1995).		
DR	EMBL: U31993; AAC52357.1; -.		
DR	MGI: 107399; 1117r.		
KW	Receptor.		
SO	SEQUENCE	864 AA; 97807 MW; 343F51AA687DA31 CRC64;	

	Query Match	4.43;	Score 165;	DB 11;	Length 864;	
	Best Local Similarity	19.8%;	Pred. No. 4.8e-05;			
	Matches 123;	Conservative	79;	Mismatches 209;	Indels 210;	Gaps 31.
Oy	192 PALPW-----LWNSAGDWNHVLVN-----VSEQHGHLSLYMMNOVGPP--KBRW--HKNL 239	: : : :				
Dd	14 PALSGLLLLLLVNLAIGRASPRLLDRPARVCABE--GLSC---RVKNSTCLDDSWIHPKML 68					
Oy	240 T--GPOITLN-----HTDLVPCLCIGVWPBEPDS-----VRTN--IC- 273	: : : : : : : : : :				
Dd	69 TPSSSKNIYINLSVSTSGELVPLVHVE-WTLQTDASILYLEGAELSVLOLNTNERLCV 127	: : : : : : : : : :				
Oy	274 --PFREDPRAHONLMQAAARLLRTLIQSMLLDAPCSLPARA-----ALCMRAPGGDP-- 322	: : : : : : : : : :				
Dd	128 KFOFLSMLOHHKKRR-----FSESHPFVD-----PGGEYEVTYHNLRPKPIRDPDNHK 176	: : : : : : : : : :				
Oy	323 -----CQPLVPLPSMENTVVDKLEFRLLKGHPNLCCVOVNSSKIOJECIADSLG 374	: : : : : : : : : :				
Dd	177 SKIIIVPROCEDSKMKMTTISCVSSSGLMD-----PNITVFELIDQHLRVDTLWNES? 228	: : : : : : : : : :				
Oy	375 PLKDVLLELTRGPODNRSLCALEPSGCTSL-----PSKASTHAARIGEYLILLDLSQG 428	: : : : : : : : : :				
Dd	229 PYQ---VLLESFSDESNHS-----CPDVVKQIFARPQEERHQRANVTFLSKFH--- 274	: : : : : : : : : :				
Oy	429 CILQMDDDLGLAMAC-----CHHHHNYOVQPFESSCLNDCLRHAATVPRCVTSNTIYVPRVADYI--PL 321	: : : : : : : : : :				
Dd	275 -----WCCHHHNYOVQPFESSCLNDCLRHAATVPRCVTSNTIYVPRVADYI--PL 321	: : : : : : : : : :				
Oy	453 WA--LVMYLACLFFAALSLILLLK-----KDHAKKAARGRA 486	: : : : : : : : : :				
Dd	322 WYGLLTLLAILLVGSVIYLLICMWRISLGADQEKHGDSKINGILIPVADLTGRPLLPRPK 381	: : : : : : : : : :				
Oy	487 ALLYSADSDSGFERIYVALSAL--QLPLRYAVDWLMSRELSAQGVAFNAHQOROTIQQ 545	: : : : : : : : : :				
Dd	382 WMIVYSADHRLPYEVYVLFKAQFLITACGTGVAALLDLBEQYISEVGMIVWSXQKQEMES 441	: : : : : : : : : :				
Oy	546 GGVVVLLFPFG-----AVAL-CSEMILQDGVSGPCAHGPHDAFNASLSCVLP 590	: : : : : : : : : :				

```

Db 442 NSKIIILCSRGTOAKWAILGWAEPVQLRCDHWKAG-----DLFTAMNWILP 491
Qy 591 DFLOGRAPSYYGACFDRLLHDPDAVPAIFRTVPVFTLPSQLPDLGALQOQPRAPRSGRLQ 650
Db 492 DFKRPACFTGVVYVVCYFGSICSDRVDLENITSRYPLMDRFEVYFIQIDLENFPCGRMH 551
Qy 651 ERAEQVSRAIQALDSYFHP 671
Db 552 HVRELTG-----DNYLQSP 565

RESULT 6
Q30764 ID O30764 PRELIMINARY; PRT; 4340 AA.
AC 030764;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE POLYKETIDE SYNTHASE MODULES 1 AND 2.
GN NIDAL.
OS Streptomyces caelestis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae.
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=36816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL-2821;
RX MEDLINE=98053867; PubMed=9393718;
RA Kakavas S.J., Katz L., Stassi D.;
RT "Identification and characterization of the niddamycin polyketide
  synthase genes from Streptomyces caelestis.";
RL J. Bacteriol. 179:7515-7522(1997).
CC -I- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
  (SDR) FAMILY
DR EMBL; AF016585; AAC46024.1;
DR InterPro; IPR002106; AA_trna_ligase-II.
DR InterPro; IPR001227; Acyltransf_domain.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Phosphopant_attach.
DR Pfam; PF00698; Acyl_transf; 3.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00109; ketoacyl-synt; 3.
DR Pfam; PF02801; ketoacyl-synt_C; 3.
DR Pfam; PF00550; pp-binding; 3.
DR PROSITE; PS00339; AA-trna_ligase_II_2; UNKNOWN_1.
DR PROSITE; PS00075; ACP_DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
KW Oxidoreductase; Phosphopantetheine; Transferase.
SQ SEQUENCE 4340 AA; 457589 MW; D59A734CB5FB795D CRC64;

Query Match 3.5%; Score 131.5; DB 2; Length 4340;
Best Local Similarity 21.5%; Pred. No. 0.25;
Matches 146; Conservative 60; Mismatches 225; Indels 247; Gaps 30;

Qy 168 GSEVRYMSTPYRYEKLNHTQQLPALPMLNVSADGDNVHLVNVSEEQHGLSLYNQV 227
Db 1692 GPRTALETAQHLQOQIRHT-----WLKVS-----HAFHSALMDPM 1728
Qy 228 QG-----PPKPRWKNLTGPOITLTHDVLVPLCICQVWPLE-----PDS 267
Db 1729 LGAFRDLTNTNYQPPTIPLISNLTG-QIADPNH-----LCTPDYWDHARHTVRFADA 1781
Qy 268 VRTNICPREDPRAHQNLQWQARLLTLQSWLLDAPCSLPA-----EALCWR 316
Db 1782 VQT-----AHQORTTYLEIGAHPTLTLLHHTLNDNPTIPLHREHPETLTALATL 1836
Qy 317 APGD-----PCQPLVPPLSW-ENVTVKVLFFLLKGLKHPNLCVQVNS 359
Db 1837 HTTGHTTTTHTTSPQTHHLDLPTVPFORDRYWMEPVRAQV-----SQP-----CA 1883

```

```

Qy 360 EKLOLQ---ECLMADSLGPKDDVLLLTETROPQNRSLCALE-----P 399
Db 1884 DRLRYRVWVEAAPEDHTVPSAERWLLILGGSDPADAGILVGEQVHGAAVQRLDIPAP 1943
Qy 400 SG-----CTSLPSKAST-----RAARLGE--YLLQDL--QSGCLOLWDDDLGAL 440
Db 1944 AGRRLLPSCSELPLVRGSPSHTEGFCGSLRSPRVAEAVALLQALGDAGADAPLWATRGAV 2003
Qy 441 WACPMDKYI---HKRWALVWLACLLFAAALSLIILLKKDHAKAAARGRAALLLYSADDSG 497
Db 2004 AAQAGEAPSVGGAGLWGLGVAGLEADRWGGLVDLPADPAPALRGARLVVSNAPDN- 2052
Qy 498 FERLVGALASALCOLPLRVAVDLMSRRLSQAQ--PVAMFHAQRRTLTQEGGVVVVLLSPG 556
Db 2063 -----QVAIR-ASGVFVRVVPAPGRPV----- 2085
Qy 557 AVALCSEWLQDG-----VSGGAGHPDAFRASLSCLVDFLOGRAPGSYVGCACFDRL--- 609
Db 2086 -----NDWAPSGTVLITGGTGALGSQVARRLALAGAPHLLLAGRRGNSGAAELVDELTA 2140
Qy 610 -----LHPDAVPALFRTVPVF-----TLPSQLPDLGALQOQR 642
Db 2141 GAETVVAACDAARDLVLATIPERPLTAVLHAACVLDGDLGTLTPRIDAVIRAK 2200
Qy 643 APRSGRLQRAEQVS-----RALQPALDSYFHPEGTPAPGRGV 680
Db 2201 ATAARHLDLTDLDLDAFLVLFSSIVGWGNGGQATYAANAALDALAH--GRRARGORA 2258
Qy 681 -----GPGAGPG--AGDG 691
Db 2259 TSIAWGPWAGSGMAAGDG 2276

RESULT 7
Q92529 ID Q92529 PRELIMINARY; PRT; 1366 AA.
AC Q92529;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE PUTATIVE LARGE PRO/ALA/GLY-RICH PROTEIN.
GN SC9F2.06C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae.
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapalite D., Eichner A., Cullum J.,
  Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
  the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL035559; CAB37473.1;
SQ SEQUENCE 1366 AA; 141042 MW; FB954569D87D029B CRC64;

Query Match 3.5%; Score 130.5; DB 2; Length 1366;
Best Local Similarity 22.4%; Pred. No. 0.065;
Matches 175; Conservative 57; Mismatches 238; Indels 311; Gaps 42;

```

QY 21 LERLVGPDATHCSPLSCLRWSDILCLPGDIVPARGPVLPAPTHLOTELVLRCQKETDC 80
 Db 675 LDGALADPDAGH---PLTIRLSEVRAALPGP--PAPVPVTRDAVFTAYLDL----- 721
 QY 81 DLCLRVAVHLAVGHWEPEDEEKEGAA-----DSGVPEPRNASL-QAQQVLSFOA 131
 Db 722 -MCLRVATRLA-----DENGIRGTAVRRLAAKVSQVHEAARRSLGPGGGLDRES 771
 QY 132 YPT-----ARCYLE-VQVPAAL-----VQFGQSVG 156
 Db 772 FETLPGCPAPARLGGTGWAPVLAEGLFVPTGSGYRFAHEELADWIGTHLDLGEALR 831
 QY 157 SVVYDCEAALGSEVRINSYTPRYEKEKLNHTQQLPALPWLNVSDGDNVHLVNSEQ 216
 Db 832 ALVHR-RDTPLTG-----HTRTLP-VPHHRI---GSVEALLLLA-RQ 868
 QY 217 HFGLSLYWNQVQGPCKPRWHKNTLGPQITLNTDILVPCICIQVWPLEDPSVRTNICPFR 276
 Db 869 H-----GVPO-----LALTLELVHAL-----D 886
 QY 277 EDPRAHQNLMOAARLRLTOSWLDAPCSLPAEAL-----CWAPGDPQCPPLVPP 329
 Db 887 RDPHS---WNAARL---LAEALTRVPDATPTDVLRLADGIAERAGDGPQTPQVFGP 938
 QY 330 LSWENVTVDKVLEPPLKGGHNLQVQVNSSEKIQLOECLWADSLGPKDDVLLLETRGPQ 389
 Db 939 AFWTAPRVPAAATRLDLR-----RLVLAD-----GPP 965
 QY 390 DNRSLCALEPGCTSLPSKASTRAARLGEYLLODQSGOCLOL-WDDDLGALWACPMKY 448
 Db 966 H-----PHLDTAG-----LLVADPTVQPLLVRFWDDERPLPATP----- 1005
 QY 449 IHKRWALVWLACLLFAAALSLLKKAARAAAL-----LLYSADD 495
 Db 1006 -----HATVATAQAALLHTRHRLDGLTEVLVDSTH 1037
 QY 496 SGFERLVGALA-----SALCQLPLRVAVDLWSRRELSAGPVAMFIAQRR-----OT 542
 Db 1038 RRADELLAVLAESEFSAALCR-----AVERWARDERPARHRAAVTHGLRTAPHARSADRT 1092
 QY 543 LOEGGVVLLFSP-----GAVALCSEWLQDGVSGPGAGHPH-----DAFRA----- 583
 Db 1093 LLRHAALVLLAGSDSPRGGAIAL-----VQDPCRDRHLPAALDLFAACDPYLP 1145
 QY 584 -SLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFR-----TVPFTLPSQLPDFLGAL 638
 Db 1146 SAVAAALPTH-----PEPVLEAFRAILLGPDAGEALRLADATTALT--HRVAALVGR 1198
 QY 639 QOPAPRSGRLOERAEQVSRL--OPALDSYFHPGTT-----PAPGRGVGPGAGPGAGD 690
 Db 1199 VTERPETAGHL---AAYVDRRLDRDPAPRAVLLPLVTRLLDDGPEPARAA--LAGVLAAD 1253
 QY 691 G 691
 Db 1254 G 1254
 RESULT 8
 Q96M86 PRELIMINARY; PRT: 1021 AA.
 ID Q96M86
 AC Q96M86;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CNA FLJ32752 FIS, CLONE TEST12001661, WEAKLY SIMILAR TO DYNEIN BETA
 DE CHAIN, FLAGELLAR OUTER ARM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

TISSUE-TESTIS:
 RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
 RA Hotuta T., Hiraoka K., Murakawa K., Takiguchi S., Kusano J.,
 RA Watanabe M., Fujinori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.,
 RT "NEDO human cDNA sequencing project."
 RL Submitt (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK057314; BAB71423.1;
 SQ SEQUENCE 1021 AA; 112942 MW; A1207BD978EA4D0D CRC64;
 Query Match 3.1%; Score 116; DB 4; Length 1021;
 Best Local Similarity 24.0%; Pred. No. 0.72;
 Matches 183; Conservative 75; Mismatches 276; Indels 228; Gaps 44;
 QY 55 PAPGPVLPAPTHLOTELVLRCQKETDCDCLRVAVHLAVGHWEPEDEEKEFGGAADSGVE 114
 Db 294 PAPGPPEPLSLKQLKILWVLRPE---CLAGALADFTTSLGRLPDENTY---APTMPF 347
 QY 115 EPRNASIQAOVLSFQAYPTARC-VLLEVOVPAALVQFGSGSVVYDCEAALGSE--- 170
 Db 348 KHSQATQPMILLPPPGHPSATLHPLTVIOKLAAYOQOQKQLOVI-----LW 401
 QY 171 ---VRIWSYTPRYEKEKLNHTQQLPALPWLNVSDGDNVHLVNSEEHQFGLSLYNNQV 227
 Db 402 PVSVVYSTLSQAYE---GH-----WLVL-----DNCHLM----- 428
 QY 228 QGPKPRWHKNTLGPQITLNTDILVPCICIQVWPLEDPSVRTNICFREDPRAHQNLW- 286
 Db 429 ---PHWPKELLQLLELGRKAVVADLESEQLLDQPE---RNVSTVHRDFR---LWL 477
 QY 287 ---QAARLRLTLOS---W-----LLDAPCSLPAEALCRRAPGDPQCPV 327
 Db 478 IVPASASLPAVLTOHSMVPVFNQSLGELGHVLID---SVELAOOVLYHQP---PTOAL- 530
 QY 328 PLSWENVTVDKVLEPPLKGGHNLQVQVNSSEKIQLOECLWADSL----- 373
 Db 531 PLLLLHGLLHRLQYTRLOAHGRWSQVTLTVQVLOTQDLWA-SLSNPRAMQELAAVS 589
 QY 374 ---GPLKD---DVLLLETRG-----PODNRSLCA-LEP-----SGCTS 404
 Db 590 FYGGLGDTEDREALISLTQACLSPPSGSVVQPHTPQSLLATLMLPELRELDMAECKA 649
 QY 405 ---LPSKASTRAARLGE-----YLLQDQSGOCLOLWDDDLGALWACPMKY 448
 Db 650 QMHLLSPPEPRLCGLSEGQPAWLLRRQSRALLSALQSRSPVWVPESSRGAQLA---ERR 706
 QY 449 IHKRWALVWLACLLFAAALSLLKKAHAKA-----AARGRAALLYSADDSGFER 500
 Db 707 LROR--LVQVNRRLLESLODLTHVIRQDESDAPWSVLGPNARRPLEGLVLETEALE--LSQ 762
 QY 501 LVGALASAL-CQL-----PLR---VAVDLWSRR-----ELSAQGVAMFHA 537
 Db 763 LVGTLQRDLDCLLQQLKGAPPCPSRRCAVAHALWTGRLPLPNRPHAPAGQPQPHWL-- 820
 QY 538 QRRQTLOEGGVVLLFSPGAVALCSEWLQDGVSGPGAGHPH-DAFRASLSCVLPDFLOGR 596
 Db 821 --RQLSRRGOLLVRYLGVA-----DASSDVPVRFVHLSAFRHRPRLLLA--LRGE 867
 QY 597 A-----PGSVYGACFDRL---LHPDAVPALFRIV--PVFTLPQSQ--LPDFLGA 637
 Db 868 AALDQNVPSNFNFGSR--GSVSSQLQYKRLMNSNPLFRVENGPNPTVPERGULLIGLV 926
 QY 638 LQOPAPRSGRLOERAEQVSRLQALPDLSYFHPGTT---PAP 676
 Db 927 LHAENPDIALQDSSPSQSPPLPPVIS-TQAPGTSDLFAP 967
 RESULT 9

```

Q9EYF6
ID Q9EYF6 PRELIMINARY; PRT; 745 AA.
AC Q9EYF6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 82.6 KDA PROTEIN.
GN YEHQ.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=O157:H7;
RX MEDLINE=20564182; PubMed=11111050;
RA Yokoyama K., Makino K., Kubota Y., Watanabe M., Kimura S.,
RA Yutsudo C.H., Kurokawa K., Ishii K., Hattori M., Abe H., Iida T.,
RA Yamamoto K., Hayashi T., Yasunaga T., Honda T., Sakakawa C.,
RA Shinagawa H.;
RT "Complete nucleotide sequence of the defective Sakai-VT1 prophage
RT carrying the shiga toxin 1 genes of the enterohemorrhagic Escherichia
RT coli O157:H7 derived from the Sakai outbreak.";
RL Gene 258:127-139(2000).
DR EMBL; AP000400; BAB19549.1; -.
SQ HYPOTHETICAL protein.
KW SEQUENCE 745 AA; 82619 MW; 37EC82409B811F43 CRC64;

Query Match 3.1%; Score 114.5; DB 2; Length 745;
Best Local Similarity 21.6%; Pred. No. 0.63;
Matches 163; Conservative 69; Mismatches 236; Indels 287; Gaps 43;

QY 21 LERLVGPDAH-----CSPGLSCLWSDILCLPGDIVPAGPVLATHLOTELVLRC 74
Db 110 LEELATLPDARKRAQVLVAKGITELF-----CAPGEIPSAKLP-MSDVFRYSRISRF 163
QY 75 OKETDC-----DLCLRVAVHLAVHGHWEPEDEKPF-----GGAADSGVEEPRNA 119
Db 164 AR-CDICIEGLCEHV--LAVQAFVEAKTQQAEEFTHLQMRSEHVTSSDDPFASEGNA 220
QY 120 SLQ-----AQVLSFOA-----YPTARCVLLEVVQVPAALVQFGSGSVV 159
Db 221 CRQYVQQLSQALWLGGISOPLIHYEAAFSQAQAACRNR--WVSESLRQLRASV--- 274
QY 160 YDCFEAA-----LGSEVRIWSTQPYREKELNHTQQL-----PALPWLNVSDG-- 203
Db 275 -DAFHARASHYHAGECLRLQALNSR-----LNCAQEMARRDSVGEVPPVPTVVGSGIA 329
QY 204 -----DNVHLV-LNV---SEEQHGLSLVWVQVQPKPRWKNLTPQIITLNLHTDLVP 254
Db 330 GEAKLDHLRLVSLGMRQWDIEHYGLRIWF-----TDPDTGSILHLS--- 371
QY 255 CLCIQVWPLEPDSVRTNICPFREDPAHONLWQAAARLRLTLQSWLL-----DAPCS 306
Db 372 -----RSWRSEQENSPAATRRFLFSQAGALAGGQIVSQAARKS 410
QY 307 LPAAALCWAPGDPCCQLVP--PLSWENVTVDKVLEPPL-----LKGHPNLC 353
Db 411 ADGELLATR-----NRLSSVPLSPDAW-----RMLSAPLRQPGIVALEYLRQRPSC 460
QY 354 V-----QVNSSEKQLQEQEL---WADSLGPKDDVLLLETRGPDNRSLCALPESGCTSLP 406
Db 461 IRPLNQVDNLFLTPVAECISLQWDSRQTL--DAQVISGE-EDN--LTLT-----SLP 509
QY 407 SKASTRAA--RIGEVLLQDLSQGCGL-----OLWDDD----- 436
Db 510 ASAPYAVERMAA--LLQQTDDPVCLVSGFVSFVQGLTLEPQVMVMTKTRAWALDAETAP 568
QY 437 ----LGALWACPMKYIHKRWALWLAACLLFAAALSLLILLKKD--HAKAAAGRAALLL 490
Db 569 VVASLPASVLPVPSTAHO-----LLMRCQALLIQLLHNGRWRYQPSAIGQAEILIA 619

```

```

QY 491 YSADDSGFERLVGALASALCOLPLRVAVDLMSRRELSAQGPVAFPHAOROTLOEGGVV 550
Db 620 NDLTAVGFYRLAHVL-----GOFRTNTESEARVEA 648
QY 551 LLFSPGVALCSEWL-----QGVSGPGAHPDAFRASLSCVLDP-----FLQGR 596
Db 649 M-----NNGVLICEQLFPLLOQOGLNRPGF--PGEFCICELR--LPDHRFRWKNHKLFL--- 698
QY 597 APGSYVGAFCEDRLHLPDAVPALFRTVPVFTLPSQL 631
Db 699 -----LLPEVGPAFPAIVDCYTSPTTL 721

RESULT 10
Q9HZE5
ID Q9HZE5 PRELIMINARY; PRT; 1193 AA.
AC Q9HZE5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE HYPOTHETICAL PROTEIN PA3063.
GN PA3063.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004730; AG06451.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1193 AA; 135097 MW; 1FE9553F34D3C191 CRC64;

Query Match 3.1%; Score 114.5; DB 16; Length 1193;
Best Local Similarity 21.7%; Pred. No. 1.2;
Matches 154; Conservative 53; Mismatches 240; Indels 263; Gaps 38;

QY 81 DCLIRVAVHLAVHGHWE-----EPEDBEKFGGAADSGV---EPRNASLQAVV--- 126
Db 293 DLRLNMAAG-RVERAWEVGOELLALRPEDRTLLADLARLCEWTGNGPRAIGFWKQLLAGA 351
QY 127 -----LSQAYPTARCVLLEVVQVPAALVQFGSGSVVYDCEAALGSEVRIW- 174
Db 352 DDPALREHAWRLSLQMFDFDSATELLAPIGAQRQMTDELDALVSHETRGTPPEGEAWL 411
QY 175 -SYTOPYEKE-----LNHTQQLPALP--WLVNSADGDNVHLVNLVNSEQHGLS 221
Db 412 RGTVQ-RYPKORLAWORLQOILEHTQOQLEETGVWARMAR-----HPFLSVKERMQWA-E 464
QY 222 LYNNQVQPKPRWKNLTPQIITLNLHTDLVPCLCIQVWPLEPD-----SV 268
Db 465 THNLFED--PQAW-KVLGVDTFAIREPEFWRLRAALAWALEODDDARAAYERMLALDI 521
QY 269 RTN-----ICPFRED--PRAHONL---WQAAAR-----LRLTLQSWLLDAPC--SLPA 309
Db 522 RLNSRDEDLIALYRDSNPKQALQVLIGSMQSRDRPRRLASALQLAENLHDWPALKSILA 581
QY 310 EAALCWAPGDPCCQLVPPPLSW-----ENVTVDKVLLEPPLKGHNLIC 353
Db 582 EAEGLPDAQS-----PYVWARLARLAEQEGHGDVAERLYREALVRFP----- 624
QY 354 VQVNSSEKQLQEQELW-----ADSLGP-----LKDDV-----LLETRGP 388

```

```
Db 625 -----GENLVRERLLWFYIDRGRRDLSAPLLAQWHGLALRDSTLWLPFASASLLE----- 675
Qy 389 QDNRSIC-----ALEPSGCTSLPSKASTRAARLGEYLLQDLSG 427
Db 676 RNDQALAWFLYKSNPNDLVQAYADALDASG-----YQDKALRLRLRLRL----- 725
Qy 428 QCLQWDDDLGALWACPMKYTHKRWALWLCALLFAAALSLL-----LLKKDHAKAAA 482
Db 726 -----DREAVRATPDS-----PATYRLRLVAVAGPQLLAQGEARRAW 761
Qy 483 RGRALLYSADDSGFERLIVGALASALCOLPRVAVDLWSRRRELSAQGPVAVPFAORROT 542
Db 762 NGEPAWL-----QLWFEFLDQLA-ATNQEPK-----DNW-----LAWARGRLKI 802
Qy 543 LOEGGVVLLSPGVAVALCEWLQGVSGPG-----AHGPHDAFRASLSCVLPDFL 593
Db 803 GRNEIOAALRSQNRAL-ORLLERGELDPQORVEALVRLHG----- 844
Qy 594 QGRAPGSYVACFD-----RLHHPDAVPALFRVPTLPSQLPDFLG 636
Db 845 -GEALGEALGALGDGHSRDNREQLRQAAILERTPGQLGWNKRDFGG 893

RESULT 11
Q96DW2
ID Q96DW2 PRELIMINARY; PRT; 657 AA.
AC Q96DW2;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SIMILAR TO RECO PROTEIN-LIKE 4 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. LYMPHOMA;
RC TISSUE-LYMPH, AND LYMPHOMA;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013277; AAH13277.1; -.
FT NON_TER 1
SQ SEQUENCE 657 AA; 73415 MW; CF51B910F560CA18 CRC64;

Query Match 3.0%; Score 112.5; DB 4; Length 657;
Best Local Similarity 19.6%; Pred. No. 0.78;
Matches 115; Conservative 73; Mismatches 194; Indels 205; Gaps 31;

Qy 16 PVLSLERLVGPQDATHCSPLGSC-----RLWDSIL--CLPGDIVPAPGPVLAPTHQ--- 67
Db 68 PCYLRVCKVLRERMGVHCFGLGTATATRTASDVAQHLLAVAEEDPLHGPAPVPTNLHLSV 127
Qy 68 -----TELVRCKQKTDCLRLVAVHL--AVHGHWEPEDEE 103
Db 128 SMRDTDQALLTLQGRFQNLDSIIYCNRRDTE---RIAALLRTCLHAAVPGS--- 181
Qy 104 KFGGAADSGVEEPNRL-----QAQVLSFQAY-----PTARCVLLEV 142
Db 182 ---GGRAKPTTAAAYHAGMCSRERRRVRQAFMVGQQLRVVATVAFGMGLDRPDVRAV--LHL 238
Qy 143 QVPAALVFGQSGSVVYDCFEAALGSEVRINWSTQPRYE--KELNHTQQLPALPWLNV 200
Db 239 GLPPSFESYVQAVGRAGRD-----GQPAHCHFLFQPGEDLRLRRHVHADSTDFLAVK 292
Qy 201 ADGDNVHLVNLV-----SEQHFGLSLYWNQVQGP-PKPRW-----HKN 238
Db 293 -----RLVQRVFPACTCTCTRPPSEQE-----GAVGERPVPKYPPQAEQLSHQA 338
Qy 143 QVPAALVFGQSGSVVYDCFEAALGSEVRINWSTQPRYE--KELNHTQQLPALPWLNV 200
Db 239 GLPPSFESYVQAVGRAGRD-----GQPAHCHFLFQPGEDLRLRRHVHADSTDFLAVK 292
Qy 201 ADGDNVHLVNLV-----SEQHFGLSLYWNQVQGP-PKPRW-----HKN 238
Db 293 -----RLVQRVFPACTCTCTRPPSEQE-----GAVGERPVPKYPPQAEQLSHQA 338
Qy 239 LTGPQIITLNLHTDLVPC-LCIQWVPLEPDSVTRNICPFREDPRAHONLWQARLRL--T 295
Db 339 APGPRVCMGHERALPIQLTQALDMPAEIETLLCYL-----ELPHHW-----LELLATT 390
```

```
Qy 296 LQSWLLDAPCSLPAAALCWRAPE-----GQDPCQ-----PLVPPPLSWENV 335
Db 391 YTHCLRLNCPGPAQLQALAHRCPPPLAVCLAQOLPEDGQSSSVEFDMVKLVDSMGWELA 450
Qy 336 TVDKVL-----EFPLKLGHPNLCVQVNSSEKQLQOECLWADSLG 374
Db 451 SVRALCOLQWHDHPTGVRGRTGVIVFEFSELAFHLRSPGDLTAEEKDOICDFLY---G 506
Qy 375 PLKDDVILLLETRGPDNRSICAL-----PSGCTSLPSKASTRAAR--LGEY 419
Db 507 RVQ-----AREQALARLRTFOAFHSVAFPPSCGCGLEQOEDEERSTRLLDLGRY 556
Qy 420 L-----LQDLOS---CQC-LQLWDDDLGALWACPMKYTHKR 452
Db 557 FEEEGEPGGMEDAQGPQARLQDMEDQV---RCDIROFLSLR 599

RESULT 12
Q96F55
ID Q96F55 PRELIMINARY; PRT; 744 AA.
AC Q96F55;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SIMILAR TO RECO PROTEIN-LIKE 4 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA, AND CHORIOCARCINOMA;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011602; AAH11602.1; -.
FT NON_TER 1
SQ SEQUENCE 744 AA; 82839 MW; 6AE0F90340A2C9F1 CRC64;

Query Match 3.0%; Score 112.5; DB 4; Length 744;
Best Local Similarity 19.6%; Pred. No. 0.92;
Matches 115; Conservative 73; Mismatches 194; Indels 205; Gaps 31;

Qy 16 PVLSLERLVGPQDATHCSPLGSC-----RLWDSIL--CLPGDIVPAPGPVLAPTHQ--- 67
Db 155 PCYLRVCKVLRERMGVHCFGLGTATATRTASDVAQHLLAVAEEDPLHGPAPVPTNLHLSV 214
Qy 68 -----TELVRCKQKTDCLRLVAVHL--AVHGHWEPEDEE 103
Db 215 SMRDTDQALLTLQGRFQNLDSIIYCNRRDTE---RIAALLRTCLHAAVPGS--- 268
Qy 104 KFGGAADSGVEEPNRL-----QAQVLSFQAY-----PTARCVLLEV 142
Db 269 ---GGRAKPTTAAAYHAGMCSRERRRVRQAFMVGQQLRVVATVAFGMGLDRPDVRAV--LHL 325
Qy 143 QVPAALVFGQSGSVVYDCFEAALGSEVRINWSTQPRYE--KELNHTQQLPALPWLNV 200
Db 326 GLPPSFESYVQAVGRAGRD-----GQPAHCHFLFQPGEDLRLRRHVHADSTDFLAVK 379
Qy 201 ADGDNVHLVNLV-----SEQHFGLSLYWNQVQGP-PKPRW-----HKN 238
Db 380 -----RLVQRVFPACTCTCTRPPSEQE-----GAVGERPVPKYPPQAEQLSHQA 425
Qy 239 LTGPQIITLNLHTDLVPC-LCIQWVPLEPDSVTRNICPFREDPRAHONLWQARLRL--T 295
Db 426 APGPRVCMGHERALPIQLTQALDMPAEIETLLCYL-----ELPHHW-----LELLATT 477
Qy 296 LQSWLLDAPCSLPAAALCWRAPE-----GQDPCQ-----PLVPPPLSWENV 335
Db 478 YTHCLRLNCPGPAQLQALAHRCPPPLAVCLAQOLPEDGQSSSVEFDMVKLVDSMGWELA 537
Qy 336 TVDKVL-----EFPLKLGHPNLCVQVNSSEKQLQOECLWADSLG 374
```

```

Db 538 SVRRALCQLQMDHEPTGVRGTVGVLFSELAFLHRLSPGDLTAEEKDQCIDFLY-----G 593
Qy 375 PLKDDVLLLETRGPQDNRSKALE-----PSGCTSLPSKASTRAAR-----LGEY 419
Db 594 RVQ-----AREQALARLRTFOAFHSVAFPSGCGCLEQOQDEERSTRLLKDLGRY 643
Qy 420 L-----LQDLQS---GOC-LQLWDDDLGALWACPMKDYIHKR 452
Db 644 FEEEGEGPGMEDAQPEPGQARLQDWEQV-----RCDIRQFLSLR 686

RESULT 13
Q9RY89
ID Q9RY89 PRELIMINARY; PRT; 478 AA.
AC Q9RY89;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2001 (TREMBlrel. 16, Last annotation update)
DE HYPOTHETICAL 51.2 KDA PROTEIN.
GN DR0061.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN STRAIN=R1;
RP SEQUENCE FROM N.A.
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vanathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001869; AAF09658.1;
DR TIGR; DR0061;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 478 AA; 51231 MW; 5E74CDEB8FBA5C95 CRC64;

Query Match 3.0%; Score 111.5; DB 16; Length 478;
Best Local Similarity 25.7%; Pred. No. 0.62;
Matches 105; Conservative 39; Mismatches 145; Indels 119; Gaps 20;

Qy 276 REDPRAHONLWQAARLLTLQSWLL-----DAPCSLPAAALCWRAPGGDPCPLV 327
Db 110 RADPREPLLEAAQ---RIRWLARLARAPQLPGTVPAHSVLY---AQEASALS 161
Qy 328 PPLSWENVTVKLEPPLKGNHLCVQVNSSEKLOQECWADSLGPKDKDDVLLLETRG 387
Db 162 APAAGDGLT-----LHLGGPD---RLAHWQGLRLS---WR-SLCPNQLLVQDETQS 206
Qy 388 PDNRSKALEPSSGCTSLPSKASTRAARLGEYLLQDLQSCQCLQLWDDDLGALWACPMK 447
Db 207 PVSQGQALLRP-----DLFASERQLEFLTGGQQLQVLFSG-----DY 244
Qy 448 YIHKRWLWVLAACLLFAAALSITLLKDKHAKAAARGA-ALLYSADDSGFERLVGALA 506
Db 245 VLLRRRA-----DAAQTRHL-----ARLAALGRACALLLLPAEQHGRRLARTLA 289
Qy 507 SALCQLPLRVAVDLNRSRRELSAQGPVAFHQAQRRTLOEGGVVLLFSPGAV----- 558
Db 290 RRLRGDPPR-ADDPASGQVAFATPAECLEAAARR-SLQRLPLELIGRFSPAQAHAQVAQ 347
Qy 559 -----ALCSEWLQDGVSP--GAHCPHDAFRASLSCVL----PDF 592
Db 348 ALLHLPPHYAAOVQAAEHAASAAEPDPLVREGAGPVPVLPATGHFIVFPLGPD 407
Qy 593 LQGRAPGSYVGACFDRL--LHPDAVPALFRTVPFTLPSQLPDFLGL 638

```

```

Db 408 LEVRLPG-----DRLLTLRPD-----YRAELVAVLPQGAATVGDGL 443

RESULT 14
Q66031
ID Q66031 PRELIMINARY; PRT; 1279 AA.
AC Q66031;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE IMMEDIATE EARLY GENE PRODUCT (HYPOTHETICAL 136.8 KDA
DE PROTEIN).
GN IRI.
OS Cercopithecine herpesvirus 9 (simian varicella virus), and
OS Cercopithecine herpesvirus 7.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=35246, 35245;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=Cercopithecine herpesvirus 9 (simian varicella virus);
RX MEDLINE=96434997; PubMed=8837883;
RA Gray W.L., Gusic N.J., Ek-Kommonen C., Kempson S.E., Fletcher T.M.;
RA "The inverted repeat regions of the simian varicella virus and
RA varicella-zoster virus genomes have a similar genetic organization.";
RL Virus Res. 39:181-193(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=Cercopithecine herpesvirus 7;
RX Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;
RA "Complete Sequence of the Simian Varicella Virus Genome.";
RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; U33499; AAB39965.1;
DR EMBL; AF275348; AAG27252.1;
DR EMBL; AF275348; AAG27249.1;
DR InterPro; IPR000923; Copper_blue.
DR PROSITE; PS00196; COPPER_BLUE; UNKNOWNL.
SQ SEQUENCE 1279 AA; 136763 MW; A3691FC354F9C877 CRC64;

Query Match 3.0%; Score 111; DB 12; Length 1279;
Best Local Similarity 21.0%; Pred. No. 2.5;
Matches 119; Conservative 39; Mismatches 174; Indels 234; Gaps 28;

Qy 266 DSVRTNCTPFREDPRAHONLWQAARLLTLQSWLLDAPCSLPAAALCWRAPGGDPC- 323
Db 713 DGPSTNGC-FRRVP-----AGALHTP--VPSEARRAYCAPNVIESL 751
Qy 324 --QPLVPPLSWENVTVKLEPPLKGNHLCVQVNSSEKLOQEC-SSSEKLOQEC 367
Db 752 VDEPMFPP-AWR-----PVLRFDPAAVAEIASRGHGVDRFRFPVPSVDALRRRC 800
Qy 368 LWADSLGPKDKDDVLLLETRGPQDNRSKALEPSSGCTSLPSKASTRAARLGEYLLQDLQSG 427
Db 801 AWMHQVAD-PEDVRLVIYDPLPGESLC-----GPTCEGGAS-----RDVH-- 841
Qy 428 QCLQWDDDLGALWACPMKDYIHKRWALWLAACLLFAAALSITLLKDKHAKAA----- 481
Db 842 ----WSDCRGGL-----SVVLAALSNRLCLPOTHAWAGNSGPP 876
Qy 482 ----ARGRAALLYSADDSGFERLVGALASALCQLPLR-----VAVDLWSR----- 523
Db 877 DVSQLNARG----ILLSTRDLAFAGAEYLGSRSLASAGRLLILDVAFERPRDGPALS 933
Qy 524 ----RELSAQGPVAFHQAQRRTLOEGGVVLLFSPGAV-----PGAVA----- 559
Db 934 QYHVYVVPAPAPDAQAVVWRPRA-----TEAGLARAVFASSTRFGPGSFARVETAFSNL 987
Qy 560 ----LC-----SEWLQDGVSG----- 571
Db 988 YPNEQPIRLCRGANVCYVWATRAGPRTRVPLSPRDYRQVYLPFGDFGCKDIARQAQGLAVG 1047

```


QY 572 -----PGAHPGPHDAFRASLSCVL-PDFL-OGRAPGSYVGACFDRLHHPDAVPAALFRTVP 623
Db 1048 AADFVDDAAYSHRAANRWGLGAPLPVLPGRPGS-AGP-----RPSDVPALARVFC 1100
QY 624 VFTLPSPDLFLGALQOPRAPRSGLRQRAQVSRALPALSYPHPPGTPA-----675
Db 1101 RHALPEPDEAPEIILPPVAGRSVAVFCMSEARESL-PPIPRIILWPPGSGGETAVEMG 1159
QY 676 -----PGRGVGPGAGPGAGD 690
Db 1160 DGTRLVFEGHGPREDDEGGAAPGCCD 1185

RESULT 15
Q9N015
ID Q9N015 PRELIMINARY; PRT; 757 AA.
AC Q9N015;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE BRAIN PROTEIN.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CEREBELLUM;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.,
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046632; BAB03550.1; --
SQ SEQUENCE 757 AA; 78303 MW; 5222A33F169B586B CRC64;

Query Match 3.0%; Score 110; DB 6; Length 757;
Best Local Similarity 20.9%; Pred. No. 1.5;
Matches 143; Conservative 53; Mismatches 213; Indels 274; Gaps 33;

QY 143 QVPAALVQFGSGSVYDCEFAALGSEVRINSYTPQRYEKELNHTQQLPALPWLNVSD 202
Db 21 QDPAA-----SDVGSV-----PPVEVYSQEPGAQPDALARSPPAE-----58
QY 203 GDNVHLVNLVSEEQHFGLSLYWQVQPKPRW-----HKNLTGQ-----243
Db 59 -----ELPVTPTKRAEVSNEVSSSPLPKQVDLPDAKDSFGPQMPDPAPSEAPDGGP 112
QY 244 ---IITLNHTDLVPCLCIQVWPLEPDSVR-----TNICPFREDPRAHONLWQAARL 291
Db 113 KPEIATMNGAD-----PISQVRGAVEAPGTPKPLIPGSDPGPAAN-----155
QY 292 RLLTLQSWLLDAPC-SLPAAALCWRAPGGDCQPLVPLSWENVTVDKVFEPFLKGGH 350
Db 156 -----RTESPMGALQDDAEW--PGRQSHPPAPP-----184
QY 351 NLCVQVNSSEKILQECLEWADSLGPKDDVLLLETRGPDNRSLCALEPSGCTSLPSKAS 410
Db 185 ---VOAPSTSRGLRVTQTRALG-----QPPPEPSASSMASAPA--SSPPANAT 230
QY 411 TRAARLGE-----YLLQDLQSGOCLQLWDDDDLGALWACPMCKYIHK 451
Db 231 APPLRWGPLRRVLSFWSLHVYGVGVFLLLPALLA-----LAALAAAPAG----P 276
QY 452 RWALVWLACLFLFAALSLIILKKDHAKAAARG-RAALILYSADDSGFERLYGALASALC 510
Db 277 RLALVAALVLVVASGLRSAYMLTDYPSGARLGVGRGLVLYNLP---FPLLLTALA-ALT 332
QY 511 QL-----PLR-----VAVDL---WSRRELSAQG-PVAMFHAORROT 542
Db 333 LIGLGAGLPPPLQNPPLLGLGVALVHGVGLLATDILLSTWNLNLLMQGLSCAW-----384

QY 543 LOEGGVVLLLFSPCAVALCSEWLQDGYSG-PGAHGPDAFRASLSCVLPDFLQ-----594
Db 385 ---GAAVAL-----GTCLCRRRLDDGPRGWDASPGPRLLAVALGALLASGLQAAALWL 437
QY 595 -----GRAPGSYVG-----ACFDRLLHDPDAV 615
Db 438 YPGPGRVGRFSWAWMGVFWLRLLELTWALALAAVAAAPRPPTTEHACWAKLMRL-AC 496
QY 616 PALFRTVPVFTLPSPDLFLGALQOPRAP-----RSGRLQERAFQVSRAL--QPALDSYF 668
Db 497 PA-----PS-----GKSEVPERPNNCYARPPSSVGAGSLDISKSLIRNPAENGQL 540
QY 669 HPPCTPAPGRGVGPGAGPGAGDG 691
Db 541 ATPSSGAWGSAASLGRGPQGGPG 563

Search completed: May 19, 2002, 15:19:08
Job time: 6744 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 14:10:19 ; Search time 43.62 Seconds
(without alignments)
614.258 Million cell updates/sec

Title: US-09-608-918-2

Perfect score: 3716

Sequence: 1 MPVPWFLLSLALGRSPVVL.....TPAPRGVGVGAGPGAGDGT 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	137.5	3.7	3707	1	PGBM_MOUSE
2	118.5	3.2	1711	1	PTPO_RAT
3	113.5	3.1	761	1	PQOF_KLEPN
4	112.5	3.0	1208	1	RCQ4_HUMAN
5	108.5	2.9	925	1	W707_HUMAN
6	107	2.9	825	1	IL4R_HUMAN
7	104.5	2.8	975	1	CUT1_CANFA
8	103.5	2.8	676	1	UL06_HSV1
9	102	2.7	922	1	W707_MOUSE
10	101.5	2.7	1151	1	ATC7_YEAST
11	100.5	2.7	753	1	UL06_HSV2B
12	99	2.7	955	1	CHRD_HUMAN
13	99	2.7	1011	1	M3K6_HUMAN
14	98.5	2.7	587	1	NDC2_RAT
15	98.5	2.7	741	1	TLE4_RAT
16	98.5	2.7	766	1	TLE4_HUMAN
17	98.5	2.7	948	1	CHRD_MOUSE
18	98.5	2.7	1505	1	CUT1_HUMAN
19	98.5	2.7	1524	1	ICP4_HSVB
20	98	2.6	1487	1	ICP4_HSVB
21	98	2.6	1487	1	ICP4_HSVB
22	98	2.6	2205	1	POLN_RUBVT
23	97.5	2.6	418	1	NER3_RAT
24	97.5	2.6	850	1	NRG2_HUMAN
25	97.5	2.6	4393	1	PGBM_MOUSE
26	97	2.6	2594	1	7LES_DROVI
27	96.5	2.6	392	1	GAG_BLVAV
28	95.5	2.6	505	1	CXAA_MOUSE
29	95	2.6	862	1	CUT1_RAT
30	94.5	2.5	614	1	YEHQ_ECOLI
31	94	2.5	993	1	AFSR_STRCO
32	94	2.5	1446	1	IE18_PVRKA
33	93.5	2.5	642	1	FZD1_MOUSE

075325 homo sapien
Q06418 homo sapien
P29915 paracoccus
Q9u449 homo sapien
P49000 rattus norv
Q9rca8 bacillus ha
P41846 caenorhabdi
Q99558 homo sapien
Q07954 homo sapien
Q60755 homo sapien
Q9up38 homo sapien
P53564 mus musculu

ALIGNMENTS

RESULT 1
PGBM_MOUSE
ID PGBM_MOUSE STANDARD; PRT; 3707 AA.
AC Q05793;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (Perlecan) (PLC).
GN HSPG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=92078153; PubMed=1744087;
RA Noonan D.M., Fullie A., Valente P., Cai S., Horigan E., Sasaki M.,
RA Yamada Y., Hassell J.R.;
RT "The complete sequence of perlecan, a basement membrane heparan
sulfate proteoglycan, reveals extensive similarity with laminin A
chain, low density lipoprotein-receptor, and the neural cell adhesion
molecule.";
RT J. Biol. Chem. 266:22939-22947(1991).
RL [2]
SEQUENCE OF 940-1601 AND 1870-2600 FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE=89034110; PubMed=2972708;
RX Noonan D.M., Horigan E.A., Ledbetter S.R., Vogeli G., Sasaki M.,
RA Yamada Y., Hassell J.R.;
RT "Identification of cDNA clones encoding different domains of the
basement membrane heparan sulfate proteoglycan.";
RL J. Biol. Chem. 263:16379-16387(1988).
CC -!- FUNCTION: THIS PROTEIN IS AN INTEGRAL COMPONENT OF BASEMENT
MEMBRANES. IT IS RESPONSIBLE FOR THE FIXED NEGATIVE ELECTROSTATIC
CHARGE AND IS INVOLVED IN THE CHARGE-SELECTIVE ULTRAFILTRATION
PROPERTIES. IT INTERACTS WITH OTHER BASEMENT MEMBRANE COMPONENTS
SUCH AS LAMININ AND COLLAGEN TYPE IV AND SERVES AS AN ATTACHMENT
SUBSTRATE FOR CELLS.
CC -!- SUBUNIT: PURIFIED PERLECAN HAS A STRONG TENDENCY TO AGGREGATE IN
DIMERS OR STELLATE STRUCTURES.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.
CC -!- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
AND O-LINKED OLIGOSACCHARIDES.
CC -!- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.
CC -!- SIMILARITY: CONTAINS 10.5 LAMININ EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 LAMININ DOMAINS IV.
CC -!- SIMILARITY: CONTAINS 15 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SEA DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	DR	EMBL; M77174; AAA39911.1; -	FT	DOMAIN	1866	1954	IG-LIKE C2-TYPE DOMAIN 4.
CC	DR	EMBL; J04054; AAA39899.1; -	FT	DOMAIN	1955	2049	IG-LIKE C2-TYPE DOMAIN 5.
CC	DR	EMBL; J04055; AAA39912.1; -	FT	DOMAIN	2050	2148	IG-LIKE C2-TYPE DOMAIN 6.
CC	DR	HSSP; Q07954; 1CR8.	FT	DOMAIN	2149	2244	IG-LIKE C2-TYPE DOMAIN 7.
CC	DR	MGD; MGI:96257; Hsp92.	FT	DOMAIN	2245	2343	IG-LIKE C2-TYPE DOMAIN 8.
DR	DR	InterPro; IPR000561; EGF-like.	FT	DOMAIN	2344	2436	IG-LIKE C2-TYPE DOMAIN 9.
DR	DR	InterPro; IPR000742; EGF_2.	FT	DOMAIN	2437	2532	IG-LIKE C2-TYPE DOMAIN 10.
DR	DR	InterPro; IPR001438; EGF_II.	FT	DOMAIN	2533	2619	IG-LIKE C2-TYPE DOMAIN 11.
DR	DR	InterPro; IPR003006; Ig_MHC.	FT	DOMAIN	2620	2720	IG-LIKE C2-TYPE DOMAIN 12.
DR	DR	InterPro; IPR003598; Ig_C2.	FT	DOMAIN	2721	2809	IG-LIKE C2-TYPE DOMAIN 13.
DR	DR	InterPro; IPR002172; LDL_recept_A.	FT	DOMAIN	2810	2895	IG-LIKE C2-TYPE DOMAIN 14.
DR	DR	InterPro; IPR000034; Laminin_B.	FT	DOMAIN	2896	2980	IG-LIKE C2-TYPE DOMAIN 15.
DR	DR	InterPro; IPR002049; Laminin_EGF.	FT	DOMAIN	2984	3162	LAMININ G-LIKE 1.
DR	DR	InterPro; IPR001791; Laminin_G.	FT	DOMAIN	3163	3241	EGF-LIKE.
DR	DR	InterPro; IPR000082; SEA.	FT	DOMAIN	3245	3425	LAMININ G-LIKE 2.
DR	DR	Pfam; PF00047; Ig; 15.	FT	DOMAIN	3426	3705	LAMININ G-LIKE 3.
DR	DR	Pfam; PF00052; laminin_B; 3.	FT	DOMAIN	3705	3705	HEPARAN SULFATE (POTENTIAL).
DR	DR	Pfam; PF00053; laminin_EGF; 8.	FT	DOMAIN	65	67	HEPARAN SULFATE (POTENTIAL).
DR	DR	Pfam; PF00054; laminin_G; 3.	FT	DOMAIN	71	73	HEPARAN SULFATE (POTENTIAL).
DR	DR	Pfam; PF00057; ldl_recept_a; 4.	FT	DOMAIN	76	78	HEPARAN SULFATE (POTENTIAL).
DR	DR	Pfam; PF01390; SEA; 1.	FT	DOMAIN	3615	3617	MEDIATES MOTOR NEURON ATTACHMENT (POTENTIAL).
DR	DR	PRINTS; PR00010; EGFBLD.	FT	DOMAIN	199	212	BY SIMILARITY.
DR	DR	ProDom; PD003031; Laminin_B; 3.	FT	DOMAIN	206	225	BY SIMILARITY.
DR	DR	SMART; SM00180; EGF_Lam; 7.	FT	DOMAIN	219	234	BY SIMILARITY.
DR	DR	SMART; SM00001; EGF_Like; 6.	FT	DOMAIN	285	297	BY SIMILARITY.
DR	DR	SMART; SM00408; IGC2; 14.	FT	DOMAIN	292	310	BY SIMILARITY.
DR	DR	SMART; SM00192; LDLA; 4.	FT	DOMAIN	304	319	BY SIMILARITY.
DR	DR	SMART; SM00281; Lamb; 3.	FT	DOMAIN	325	337	BY SIMILARITY.
DR	DR	SMART; SM00282; LamG; 3.	FT	DOMAIN	332	350	BY SIMILARITY.
DR	DR	SMART; SM00200; SEA; 1.	FT	DOMAIN	344	359	BY SIMILARITY.
DR	DR	PROSITE; PS00022; EGF_1; 8.	FT	DOMAIN	375	394	BY SIMILARITY.
DR	DR	PROSITE; PS01186; EGF_2; 5.	FT	DOMAIN	388	403	BY SIMILARITY.
DR	DR	PROSITE; PS01248; LAMININ_TYPE_EGF; 11.	FT	DOMAIN	428	479	BY SIMILARITY.
DR	DR	PROSITE; PS50025; LAM_G_DOMAIN; 3.	FT	DOMAIN	766	780	BY SIMILARITY.
DR	DR	PROSITE; PS01209; LDLRA_1; 4.	FT	DOMAIN	795	811	BY SIMILARITY.
DR	DR	PROSITE; PS50068; LDLRA_2; 4.	FT	DOMAIN	814	829	BY SIMILARITY.
DR	DR	PROSITE; PS50024; SEA; 1.	FT	DOMAIN	816	839	BY SIMILARITY.
DR	DR	Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;	FT	DOMAIN	842	851	BY SIMILARITY.
DR	DR	Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;	FT	DOMAIN	1159	1168	BY SIMILARITY.
DR	DR	Extracellular matrix; EGF-like domain.	FT	DOMAIN	1161	1175	BY SIMILARITY.
FT	CHAIN	POTENTIAL.	FT	DOMAIN	1178	1187	BY SIMILARITY.
FT	CHAIN	BASIMENT MEMBRANE-SPECIFIC HEPARAN	FT	DOMAIN	1190	1206	BY SIMILARITY.
FT	CHAIN	SULFATE PROTEOGLYCAN CORE PROTEIN.	FT	DOMAIN	1209	1224	BY SIMILARITY.
FT	CHAIN	SEA.	FT	DOMAIN	1211	1234	BY SIMILARITY.
FT	CHAIN	LDL-RECEPTOR CLASS A 1.	FT	DOMAIN	1237	1246	BY SIMILARITY.
FT	CHAIN	LDL-RECEPTOR CLASS A 2.	FT	DOMAIN	1249	1263	BY SIMILARITY.
FT	CHAIN	LDL-RECEPTOR CLASS A 3.	FT	DOMAIN	1275	1287	BY SIMILARITY.
FT	CHAIN	LDL-RECEPTOR CLASS A 4.	FT	DOMAIN	1277	1293	BY SIMILARITY.
FT	CHAIN	IG-LIKE C2-TYPE DOMAIN 1.	FT	DOMAIN	1295	1304	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 1 (N-TERMINAL).	FT	DOMAIN	1307	1322	BY SIMILARITY.
FT	CHAIN	LAMININ DOMAIN IV 1 (DOMAIN III A).	FT	DOMAIN	1563	1572	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 1 (C-TERMINAL).	FT	DOMAIN	1565	1579	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 2.	FT	DOMAIN	1582	1591	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 3.	FT	DOMAIN	1594	1610	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 4 (INCOMPLETE).	FT	DOMAIN	1613	1628	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 5 (N-TERMINAL).	FT	DOMAIN	1615	1638	BY SIMILARITY.
FT	CHAIN	LAMININ DOMAIN IV 2 (DOMAIN III B).	FT	DOMAIN	1641	1650	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 5 (C-TERMINAL).	FT	DOMAIN	1653	1668	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 6.	FT	DOMAIN	1792	1839	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 7.	FT	DOMAIN	1886	1932	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 8.	FT	DOMAIN	1976	2021	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 9 (N-TERMINAL).	FT	DOMAIN	2073	2118	BY SIMILARITY.
FT	CHAIN	LAMININ DOMAIN IV 3 (DOMAIN III C).	FT	DOMAIN	2170	2215	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 9 (C-TERMINAL).	FT	DOMAIN	2268	2313	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 10.	FT	DOMAIN	2365	2413	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 11.	FT	DOMAIN	2456	2506	BY SIMILARITY.
FT	CHAIN	IG-LIKE C2-TYPE DOMAIN 2.	FT	DOMAIN	2554	2599	BY SIMILARITY.
FT	CHAIN	IG-LIKE C2-TYPE DOMAIN 3.	FT	DOMAIN	2641	2686	BY SIMILARITY.
FT	CHAIN		FT	DOMAIN	2831	2876	BY SIMILARITY.

RESULT	2
PTPO_RAT	
ID	PTPO_
AC	Q6461
DT	01-NO
DT	01-NO
DT	15-JU
DE	Osteo
DE	(OST-
OS	Rattus
OC	Eukar

```
FT DOMAIN 563 652 FIBRONECTIN TYPE-III 7.
FT DOMAIN 563 741 FIBRONECTIN TYPE-III 8.
FT DOMAIN 742 830 FIBRONECTIN TYPE-III 9.
FT DOMAIN 831 921 FIBRONECTIN TYPE-III 10.
FT DOMAIN 1150 1418 PROTEIN-TYROSINE PHOSPHATASE 1.
FT DOMAIN 1419 1711 PROTEIN-TYROSINE PHOSPHATASE 2.
FT ACT_SITE 1350 1350 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 117 117 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 431 431 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 570 570 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 649 649 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 663 663 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 851 851 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 882 882 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 970 970 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 982 982 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1711 AA; 187292 MW; 0F04D2D1A47A18A0 CRC64;

Query Match 3.2%; Score 118.5; DB 1; Length 1711;
Best Local Similarity 18.9%; Pred. No. 0.37;
Matches 151; Conservative 81; Mismatches 272; Indels 293; Gaps 37;

QY 5 WFLLSALG-RSPVVLRLVGPQDATHCSPGLSRLWSDILPLGDIVPAPGPVLAP 63
DB 322 WASWAGLARGDGYVL---KLSGPMESTSTLGPPECN-----AVFPGP-LPP 364

QY 64 THLOELVLRCKEYDTCDCILRVAVHLVGHWEPEDEKFGGADSGVEE---PRNA 119
DB 365 GHYTLQL-----KVLAPDYADWVEGSTWLAESA 392

QY 120 SIQAQVW-----LSFOAYTARCVLLEVQVPAALVQFGOSVGS--VYDCEFAALGS 169
DB 393 ALPREVPGARLWLDGLEASKQPGRRALLYDDAPGSLGNISVPSGATHYIFCGLVPGAHY 452

QY 170 EVRIWISYTPQRYEKELNHTHQQLP--ALPWLNVSDAGDNVHLVNLVSEEHFGLSLYNQV 227
DB 453 RVDIASSTGDISQISGYTSPPLPQSLVSRSSPSD-----LTIANGPA 497

QY 228 QGP---PKPRWHKNT-----GPOITLHTDLVP--CLCIQVWP-----LEPDSV 268
DB 498 PQLEGYKVTWHODGQSQRSGDLVDLGPDTLSLTLSKSLVPGSCYTVSAWAGNLDSDSQ 557

QY 269 RTNCP-----FREDPRAONLW-----QARLRLTLQ----- 297
DB 558 KIHSTRFPAPPTNLSLGFHAQPAALKASWYHPGGRDADFHLRLYRLPLTLESEKVLPRE 617

QY 298 ---SML-LDAPCSLPAEALCW-----RAPGDDCPQPLVPPLSWENVTVDKVLFP 344
DB 618 AQNSWAQLTACCEQVQVQLTUNGSSRSANATGWTP--PSAPTL--VNVTSADPTQLQ 673

QY 345 LKGHNPCLCVQVNSSEKQLQQLWADS-----LGPLKDDVLLILETRGPQDNRSICAL 397
DB 674 VSWAHV-----PGRSRYQVTLQESTRTATSIINGP-----KEDGTSFLGL 714

QY 398 EPS-----GCTSLPSKASTRAARLGEY-----LLQDLQSQCLQLWDDDLGALWACP 444
DB 715 TPGTKYKVEISWAGPLTAANVSAWYPLIPNELLVSMQAGSAVV----- 761

QY 445 MDKIHKRWALVWLACLLFAALSLILLKDKHAKAARGRALLIYSADDSGFERLVGA 504
DB 762 -----NLAWPSGPGQGGAC-----HAQLSDAGHLS----- 786

QY 505 LASALCQLPLRVAVDLWSRRELSAQGPVAVHFAQRROTQEQ--GVVYVLLFSPCAVALCSE 563
```

```
DB 787 -----WEQPLKQLQELFMLRDLTPGHTISMSVRCRAGPLOASTHLVLVSVBPGPV----- 836
QY 564 WLQDVGSGPAGHDPDAFRASLSCLVLPD-----FLOGRAGPS-----YVGACFDRL 609
DB 837 --EDVLCHP-----EATYALNWTMPAGDDVDCLVVVVERLVPGGTHFVFOVNTSGDAL 888
QY 610 LHDAVPALFTVPVFTLPSQLPDFLQALQPRAPRSRQLQRAEQVSRALQPALDSTFPH 669
DB 889 LLPNLMTPTSYRLSLTVL-----GRNSWSRAVSLVCSAEAW-H 928
QY 670 PPGTPAP-----GRGVG 681
DB 929 PPELAEPQVELGTGMG 945

RESULT 3
PQOF_KLEPN STANDARD; PRT; 761 AA.
AC P27508;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Coenzyme PQQ synthesis protein F (EC 3.4.99.-).
GN PQQF.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 418;
RX MEDLINE=9221293; PubMed=1313537;
RA Meulenbergh J.J.M., Sellink E., Riegman N.H., Postma P.W.;
RT "Nucleotide sequence and structure of the Klebsiella pneumoniae pqq operon.";
RL Mol. Gen. Genet. 232:284-294(1992).
CC -1- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ) BIOSYNTHESIS. IT IS THOUGHT THAT THIS PROTEIN IS A PROTEASE THAT CLEAVES PEPTIDES BOND IN A SMALL PEPTIDE (GENE PQA) WHICH WOULD PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH ARE NECESSARY FOR THE SYNTHESIS OF PQQ.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE INSULINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X58778; CAA41584.1;
CC PIR; S20458; S20458.
CC MEROPS; M16.006;
CC InterPro; IPR001431; Peptidase_M16.
CC Pfam; PF00675; Peptidase_M16; 1.
CC PROSITE; PS00143; INSULINASE; 1.
CC PQQ; Hydrolase; Metalloprotease; zinc.
CC METAL 49 49 ZINC (BY SIMILARITY).
CC ACT_SITE 52 52 BY SIMILARITY.
CC METAL 53 53 ZINC (BY SIMILARITY).
CC METAL 130 130 ZINC (BY SIMILARITY).
CC SEQUENCE 761 AA; 83616 MW; 187182DBC8E839B5 CRC64;

Query Match 3.1%; Score 113.5; DB 1; Length 761;
Best Local Similarity 23.4%; Pred. No. 0.32; Indels 167; Gaps 27;
Matches 117; Conservative 50; Mismatches 165;

QY 282 HQNLQWAAARLRLTLQSWLLDAPCSLPAEALCWRAPG---DPCQPLVPPLSWENVTVD 338
DB 182 HRTHYVARMQL-----W-LOGPQSLEALGELAARFAAGLAAGEAPPAPPL----- 227
```

```
QY 339 KYLEPPLKHPNLCVYNSSKELQLOECLWADSLGPKODVILLETGRPODNRSLCALE 398
Db 228 RUGEFTALQ-----LAVSS-----OPALWRCPLIALSONVTLLR----- 261
QY 399 PSCTSLSKASTRAARIGELYLLQD-----LQSGCQLQWDDDLGALWACPMCKYIHKRWA 454
Db 262 -----EFLDEAPGSLMAGLRQRLAGDVALNW-----LYQDRH 295
QY 455 LWWLACLLFAA-----ALSILLLKKDKHAKAARGRAALL----- 489
Db 296 LGWLA-LVFASDRPEEDVROITHWLAQLOOTTPQEQOHHYLSRRRFOALSPLDQLRORA 354
QY 490 ----LYSADDSGFERLVCAL-----ASALCO-----LPLRAVDLMSRRLS 527
Db 355 FGPAPGAPAGAFDCAALQVAPSVSLACQTSVSPGPVATQGFSLPL-----SRWRRPES 410
QY 528 AQGPVAFWHAQRQTL-----QEGGVVLLFSPG-----AVALGSEWLQDGVVS----- 570
Db 411 DPALATAFYPOAGDLVAKCPKAAAPLLHLPSGDPPLRLRPFFYCSPOQAEGLARGEQ 470
QY 571 -----GPGAHPHDAFRAS--LSCVLPDFLQGRAPSGVYGACFDRLRLHPDA--VPA 617
Db 471 LRPLLAALRHAGCHGHWLFDGWSNLTQLPE--PGRPEAILQAILQLALPVASLTPS 528
QY 618 LFTVPVFTLPSQLPDLFALQOPR-----APRSRLQERAEOVSRLQPALDSYHPHGP 674
Db 529 P-ESIAIRHLMALPRLGTSGHQKGLAALAGGSAEDQWVARQLS-LITAPVNP-PP 585
QY 675 APG---RGVGPAGGAGD 690
Db 586 APAPCRGVERLVPG-GD 603

RESULT 4
RCQ4.HUMAN
ID RCQ4.HUMAN STANDARD; PRT; 1208 AA.
AC O94761;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE ATP-dependent DNA helicase Q4 (RecQ protein-like 4).
GN RECQ4 OR RECQ4.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097344; PubMed=9878247;
RA Kitao S., Ohsugi I., Ichikawa K., Goto M., Furuichi Y., Shimamoto A.;
RT "Cloning of two new human helicase genes of the RecQ family:
RT biological significance of multiple species in higher eukaryotes.";
RL Genomics 54:443-452(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20021764; PubMed=10552928;
RA Kitao S., Lindor N.M., Shiratori M., Furuichi Y., Shimamoto A.;
RT "Rothmund-thomson syndrome responsible gene, RECQ4: genomic structure
RT and products.";
RL Genomics 61:268-276(1999).
CC -!- SUBCELLULAR LOCATION: Nuclear (potential).
CC -!- DISEASE: DEFECTS IN RECQ4 ARE A CAUSE OF ROTHMUND-THOMSON
CC SYNDROME (RTS). A DISEASE CHARACTERIZED BY DERMATOLOGICAL FEATURES
CC SUCH AS ATROPHY, PIGMENTATION, AND TELANGIECTASIA AND FREQUENTLY
CC ACCOMPANIED BY JUVENILE CATARACT, SADDLE NOSE, CONGENITAL BONE
CC DEFECTS, DISTURBANCES OF HAIR GROWTH, AND HYPOGONADISM.
CC -!- SIMILARITY: BELONGS TO THE HELICASE FAMILY. RECQ SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
Cc -----
DR EMBL; AB006532; BAA74453.1; -
DR EMBL; AB026546; BAA86899.1; -
DR MIM; 603780; -
DR MIM; 268400; -
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
KW Hydrolase; Helicase; ATP-binding; Nuclear protein.
FT NP_BIND 502 509 ATP (POTENTIAL).
FT SITE 605 608 DEAD BOX.
SQ SEQUENCE 1208 AA; 133076 MW; CB809A7765AB48A1 CRC64;

Query Match 3.0%; Score 112.5; DB 1; Length 1208;
Best Local Similarity 19.6%; Pred. No. 0.7; Indels 205; Gaps 31;
Matches 115; Conservative 73; Mismatches 73;

QY 16 PVVLSLERLVGQDATHCPSGLSC--RLWDSIL--CLPGDITVPAPGPVLAATHLQ--- 67
Db 619 PCYLRCKVLRRMGVHCFGLTATATRTASDAQHLAVAEEDPLHGPVPTNLHLSV 678
QY 68 -----TELVRCOKETDCDCLRVAVHL--AVGHWEPEDEE 103
Db 679 SMDRDTQALLTLLOGKRFQNLDSIIICNRRDTE---RIALLRTCLHAAPVGS--- 732
QY 104 KFGAADSQVEEPRNASL-----QAQVLSFQAY-----PTARCVLLEV 142
Db 733 --GGRAPKTTAEYHAGMCSRRRRVQRAFPMQGLRVVATVAFGMGLORPDVRAV-LHL 789
QY 143 QVPAALVQFGSVGVYDCFAAALGSEVRINSYTPRYE--KELNHTQQLPALPWLNV 200
Db 790 GLPPSFESVQAVGRAGRD-----GQPAHCHFLFLOQGEDLELRHVVHADSTFLAVK 843
QY 201 AGDNNVHLNV-----SEEHFGLSLYNVOGP-PKPRW-----HKN 238
Db 844 -----RLQVRVFPACTCTCTRPPEQ-----GAVGGERPVKYPQEAQLSHQA 889
QY 239 LTGPQIITLHHTDLVPC-LCIQWPLEPDSVTRNICPFREDPRAHONLQAARLRL--T 295
Db 890 APGPRRVMGHERALPIQLTVQALDMPPEAETLLCYL----ELHPHW-----LELLAT 941
QY 296 LQSWLIDAPCSLPAEALCWRAE-----GGDPQC-----PLVPLSWENV 335
Db 942 YTHCLNCPGGPAQLQALAHRCPPPLAVCLAAQQLPEDPGQSSSVFDMYKLVDSMGWELA 1001
QY 336 TVDKVL-----EPPLLKGHNLCVQVNSSEKIQLOECLWADSLG 374
Db 1002 SVRRALCOLQWDHEPRTGVRRGTGVLVESELAFLHRLSPGDLTAEEKQICDFLY----G 1057
QY 375 PLKDDVLLILETRGPQDNRSICAL-----PSGCTSLPSKASTRAAR-----LGEY 419
Db 1058 RVQ-----AREQALRLRRTFOAFHVSFAFPSCGFCLEQDDEERSTRUKDLGRY 1107
QY 420 L-----LQDLQS---GQC-LQLWDDDLGALWACPMCKYIHKR 452
Db 1108 FEEEGQEPGGMGAOGPEPGQARLQDWDQV-----RCDIRQFLSLR 1150

RESULT 5
W70T.HUMAN
ID W70T.HUMAN STANDARD; PRT; 925 AA.
AC P57737;
DT 16-OCT-2001 (Rel. 40, Created)
```

Qy	329	PLSW---ENVTVDKVLBFPLKKGPHNLCVQVNSSEKIQLOECINWADSLGPKLDDV---LLL 383
Db	369	PHGWAGDNGQVQKSLNPACRPHPS-----FTSCL-VPPAEPLPDTAQPAVM 415
Qy	384	ETR-GPDNRSCLALEPSGCTSLPSKASTRAARL-----GEYLLQDLQSGQCLOLWDDDLG 438
Db	416	ETPVGDADAGEGFSPPSSULTS-PTSPSSLGPSLSSTSGTIGTSPSURLSQSL-----LG 468
Qy	439	ALWACPMDKYIHKRWALVWLACLLFAAALSILILLKKDHAKAARAALLLYSADDSGF 498
Db	469	-----PSKFRHAQGT-----VLHRDSHITNLKC---LNLTPGSGDGF 504
Qy	499	ERLVGALASALCOLPLRVAVDLSMRRLSAGQPVAMFHAQR-----RQTLQEGGVVV 550
Db	505	-----CANKLRVAVPL-----LSSGGQVAVLELRKPGRLPDTALPTLQNGAAVT 548
Qy	551	LL-----FSPGAVCALSC 563
Db	549	DLAWDPDPHRLAVAGE 565
RESULT 6		
IL4R_HUMAN		
ID	IL4R_HUMAN	STANDARD; PRT; 825 AA.
AC	P24394;	
DT	01-MAR-1992 (Rel. 21, Created)	
DT	01-MAR-1992 (Rel. 21, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Interleukin-4 receptor alpha chain precursor (IL-4R-alpha) (CD124 antigen).	
DE	IL4R OR IL4RA OR 582J2.1.	
GN	Homo sapiens (Human).	
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxId=9606;	
RP	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=peripheral blood;	
RX	MEDLINE=90171849; PubMed=2307934;	
RA	Izderda R.L., March C.J., Mosley B., Lyman S.D., Bos T.V.,	
RA	Gimpel S.D., Din W.S., Grabstein K.H., Widmer M.B., Park L.S.,	
RA	Cosman D., Beckman M.P.;	
RT	"Human interleukin 4 receptor confers biological responsiveness and	
RT	defines a novel receptor superfamily.";	
RL	J. Exp. Med. 171:861-873(1990).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=91120547; PubMed=2278997;	
RA	Galizzi J.-P., Zuber C.E., Harada N., Gorman D.M., Djossou O.,	
RA	Kastelein R., Banchereau J., Howard M., Miyajima A.;	
RT	"Molecular cloning of a cDNA encoding the human interleukin 4	
RT	receptor.";	
RL	Int. Immunol. 2:669-675(1990).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=99425270; PubMed=10493829;	
RA	Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,	
RA	Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Crohn L.,	
RA	Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,	
RA	Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;	
RT	"Genome duplications and other features in 12 Mb of DNA sequence from	
RT	human chromosome 16p and 16q.";	
RL	Genomics 60:295-308(1999).	
RN	[4]	
RP	VARIANTS VAL-75; ALA-400; ARG-431; LEU-436 AND PRO-786.	
RX	MEDLINE=97224413; PubMed=9070874;	
RA	Deichmann K., Bardutzky J., Forster	
RT	"Common polymorphisms in the coding part of the IL4-receptor gene.";	
RL	Biochem. Biophys. Res. Commun. 231:696-697(1997).	
RN	[5]	
RP	VARIANT ATOPIC ARG-576.	
RX	MEDLINE=98041803; PubMed=9392697;	
RA	Hershey G.K.K., Friedrich M.F., Esswein L.A., Thomas M.L.,	

RA Chatila T.A.;
 RT "The association of atopy with a gain-of-function mutation in the
 RT alpha subunit of the interleukin-4 receptor.";
 RL New Engl. J. Med. 337:1720-1725(1997).
 RN [6]
 RP VARIANT ATOPIC ASTHMA VAL-75.
 RX MEDLINE=98282087; PubMed=9620765;
 RA Mitsuhashi H., Izuhara K., Mao X.-Q., Gao P.S., Arinobu Y., Enomoto T.,
 RA Kawai M., Sasaki S., Dake Y., Hamasaki N., Shirakawa T., Hopkin J.M.;
 RA "Ile50Val variant of IL4R alpha upregulates IGE synthesis and
 RT associates with atopic asthma.";
 RL Nat. Genet. 19:119-120(1998).
 RN [7]
 RP VARIANT ATOPIC ASTHMA VAL-75
 RX MEDLINE=99322293; PubMed=10390422;
 RA Noguchi E., Shibasaki M., Arinami T., Takeda K., Yokouchi Y.,
 RA Kobayashi K., Imoto N., Nakahara S., Matsui A., Hamaguchi H.;
 RA "No association between atopy/asthma and the Ile50Val polymorphism of
 RT IL-4 receptor.";
 RL Am. J. Respir. Crit. Care Med. 160:342-345(1999).
 RN [8]
 RP VARIANTS PRO-503 AND ARG-576.
 RX MEDLINE=99250314; PubMed=10233717;
 RA Kruse S., Japha T., Tedner M., Sparholt S.H., Forster J., Kuehr J.,
 RA Deichmann K.A.;
 RA "The polymorphisms S503P and Q576R in the interleukin-4 receptor alpha
 RT gene are associated with atopy and influence the signal
 RT transduction.";
 RL Immunology 96:365-371(1999).
 RN [9]
 RP VARIANT ALA-752
 RX MEDLINE=20143377; PubMed=10677312;
 RA Ober C., Leavitt S.A., Tsalenko A., Howard T.D., Hoki D.M., Daniel R.,
 RA Newman D.L., Wu X., Parry R., Lester L.A., Solway J., Blumenthal M.,
 RA King R.A., Xu J., Meyers D.A., Bleeker E.R., Cox N.J.;
 RA "Variation in the interleukin 4-receptor alpha gene confers
 RT susceptibility to asthma and atopy in ethnically diverse
 RT populations.";
 RL Am. J. Hum. Genet. 66:517-526(2000).
 RN [10]
 RP VARIANT ATOPIC ARG-576.
 RX MEDLINE=20269830; PubMed=10809862;
 RA Oiso N., Fukui K., Ishii M.;
 RA "Interleukin 4 receptor alpha chain polymorphism Gln551Arg is
 RT associated with adult atopic dermatitis in Japan.";
 RL Br. J. Dermatol. 142:1003-1006(2000).
 RN [11]
 RP VARIANT PRO-786
 RX MEDLINE=21405389; PubMed=11513543;
 RA Andrews R.P., Burrell L., Rosa-Rosa L., Cunningham C.M.,
 RA Brzezinski J.L., Bernstein J.A., Khurana Hershey G.K.;
 RA "Analysis of the Ser786Pro interleukin-4 receptor alpha allelic
 RT variant in allergic and nonallergic asthma and its functional
 RT consequences.";
 RL Clin. Immunol. 100:298-304(2001).
 RN [12]
 RP VARIANT ILE-579.
 RX MEDLINE=21182061; PubMed=11285129;
 RA Lozano F., Places L., Vila J.M., Padilla O., Arman M., Gimferrer I.,
 RA Suarez B., Lopez de la Iglesia A., Miserachs N., Vives J.;
 RA "Identification of a novel single-nucleotide polymorphism (Val554Ile)
 RT and definition of eight common alleles for human IL4RA exon 11.";
 RL Tissue Antigens 57:216-220(2001).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR INTERLEUKIN-4. A SOLUBLE FORM
 CC OF THE IL-4 RECEPTOR MAY REPRESENT A REGULATORY MOLECULE SPECIFIC
 CC FOR IL-4-DEPENDENT IMMUNE RESPONSES.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA CHAIN AND A COMMON GAMMA CHAIN.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- DISEASE: DEFECTS IN IL4R ARE A CAUSE OF ATOPIC ASTHMA.
 CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
 CC -1- DATABASE: NAME=PRO; NOTE=CD guide CD124 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd124.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X52425; CAA36672.1; -
 DR EMBL; AC004525; AAC23495.1; -
 DR PIR; A60386; A60386.
 DR PIR; A47603; A47603.
 DR MIM; 147781; -
 DR InterPro; IPR002996; CRIA.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003531; Hematopo_receptor_S_F1.
 DR Pfam; PF00041; fn3; 1.
 DR SMART; SM00060; FN3; 1.
 DR PROSITE; PS01355; HEMATOPO_REC_S_F1; 1.
 KW Receptor; transmembrane; Glycoprotein; Signal; Disease mutation;
 KW Polymorphism.
 FT SIGNAL 1 25
 FT CHAIN 26 825
 FT DOMAIN 26 232
 FT TRANSMEM 233 256
 FT DOMAIN 257 825
 FT DISULFID 34 44
 FT CARBOHYD 74 86
 FT CARBOHYD 53 53
 FT CARBOHYD 98 98
 FT CARBOHYD 128 128
 FT CARBOHYD 134 134
 FT CARBOHYD 176 176
 FT CARBOHYD 209 209
 FT VARIANT 75 75
 FT VARIANT 400 400
 FT VARIANT 431 431
 FT VARIANT 436 436
 FT VARIANT 503 503
 FT VARIANT 576 576
 FT VARIANT 579 579
 FT VARIANT 752 752
 FT VARIANT 786 786
 FT SEQUENCE 825 AA; 89658 MW; 9F86DF5612297F8 CRC64;
 SQ
 Query Match 2.9%; Score 107; DB 1; Length 825;
 Best Local Similarity 19.2%; Pred. No. 1.2;
 Matches 155; Conservative 66; Mismatches 284; Indels 304; Gaps 36;
 QY 41 LWDSDILCPGDIIVP-----APGPVLAPTHQLQELVLRCKQKDCDCLRLVAVHVLAVH 93
 Db 103 LWAGQQLLWKGSKFSEKSHVPRAPGNLTVHTNVSDTLTLLTWSNPYPDPNYLNHLTYAVN 162
 QY 94 GHWEPEDEKFGGADSGVEEPRNASLQAQVLSFQAYPTARCVLLE--VQVPAALVQF 151
 Db 163 -INSENDPAD-----FRIV---NVTLEPSLRIAASTLKS 193
 QY 152 GQSVGSVVYDCFEAALGSEVRINSYTOPRYEKELNHHQQLPALPWLNVNSADGNVHLVLN 211
 Db 194 GIS-----YRARVANAQCY-----NTTWESESPSTKWHNSREPPFQHLHLG 236
 QY 212 VSEQHFGLSL-----YWNQVGGPKPKR-----WHKNLTPQI 244

```

Db 237 VSVSCIVLAVCLLCVSVITKIKKWDQIPNPARSRLVAILIQQAGSGWERSRGQEP 296
Qy 245 ITLNH-----TDLVPLCLCIQVWPDSVRTNICPPREDPRAHQNLQWAAARLLTLQSW 299
Db 297 AKCPHKKNCLTKLLPCF-----LEHMKRD-----EDPH-----KAA--KEMPFQGS 336
Qy 300 LLDAPCSLPAEALCWRAPGDPQPLVPLSWENVYDKVLEFPL-----LKGH 349
Db 337 GKSAMCPVEISKTVLW-----PESIVVRCVVELFEAPVECEEEVEBEK- 382
Qy 350 PNLCOVQVNSSEK-----LQLECLWADSLGPKDDVLLLETRG-----PODNRSLCA 396
Db 383 -SFCASPESSRDDFOEGREGIVARITESLFLDLG-----EENGFCQODMGESCL 432
Qy 397 LEPSGCTG-----LPSKASTRAARLGEYLLQDLQSGQCLQDLWDDLLGALWACPMDKYI 449
Db 433 LPPSGSTSAHPWDEFPAGKEAPPWK-----EQPLHEPFPASPTQSPN--- 481
Qy 450 HKRWLVWLACLLFAAALLSLLLLKKHAKAARGAALLYSADDSGFERLVGALASAL 509
Db 482 -----LTC-----TETPLVIAGNPAYRFSFNSLSQSP 508
Qy 510 QCLPLRAVDLWSREL-----AQPVAWFHAQRQRIQLEGVVV 550
Db 509 C--PRELGPDPLLARHLHEVEPEMPCVPQLSEPTTVQPEPETWEQILRRNVLQHGAAA 566
Qy 551 LLFSP--GAVALCSEWLQGV-----GPCAGHPHDAFRASL--SCVLPDFLQ--GRAP 598
Db 567 PVSAPTSYQYEVHVEGQGTQASAVGLGPPGEAGYKAFSSLLASAVSPEKCGFGASS 636
Qy 599 GSYVGACDRLL-----HPDAVPALFRIVPVT-----LPSQLPDFLGA- 637
Db 627 GEEGYKPPQDLIPGCGDPAPVP-----VPLFTFGLDREPPRSPQSSHLPSSEPHLGL 681
Qy 638 -----LQOPRAPS-----GRQERAEQVSRAQLPALDS 666
Db 682 PGEKVEDMPKPLPOEQATDPLVDSLGSGIVYSALTCHLCGLKQCHGQEDGGQTPVMAS 741
Qy 667 -----YFHPPTGTPAPGRGVGPGAGP 686
Db 742 PCCGCCCGDRSPPTTPLRAPDPSGGVP 770

RESULT 7
CUT1_CANFA
ID CUT1_CANFA STANDARD; PRT; 975 AA.
AC P39881;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE CCAAT displacement protein (Homeobox protein Clox) (Clox-1)
DE (Fragment).
GN CUT1 OR CLOX.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart ventricle;
RA MEDLINE=93161940; PubMed=1363085;
RA Andres V., Nadal-Ginard B., Mahdavi V.;
RT "Clox, a mammalian homeobox gene related to Drosophila cut, encodes
RT DNA-binding regulatory proteins differentially expressed during
RT development.
RL Development 116:321-334(1992).
CC -!- FUNCTION: DNA-BINDING REGULATORY PROTEIN, WHERE THE SMALLER
CC PROTEIN SPECIES PROBABLY REPRESENTS THE ACTIVE DNA-BINDING FORM.
CC MAY FUNCTION AS A DEVELOPMENTALLY REGULATED REPRESSOR OF TISSUE-
CC SPECIFIC GENE TRANSCRIPTION BY PREVENTING THE INTERACTION OF
CC TISSUE-SPECIFIC TRANSCRIPTIONAL FACTORS WITH THEIR COGNATE TARGET
CC SEQUENCES. PROBABLY INVOLVED IN CELL-FATE SPECIFICATION IN DIVERSE

```

```

CC DIFFERENTIATION PROGRAMS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: A number of isoforms may be produced by
CC alternative splicing.
CC -!- TISSUE SPECIFICITY: A BROAD PATTERN OF EXPRESSION OBSERVED IN
CC TISSUES OF DIVERSE ORIGINS, SUCH AS CARTILAGE, LIVER, BRAIN, LUNG,
CC HEART AND SKELETAL MUSCLE. THERE ARE 2 DISTINCT PROTEIN SPECIES:
CC THE LARGER ONE (230-250 KDA) IS FOUND MAINLY IN ADULT BRAIN, LUNG
CC AND HEART, AND THE SMALLER ONE (180-190 KDA) PREDOMINATES IN EARLY
CC EMBRYONIC TISSUES.
CC -!- DEVELOPMENTAL STAGE: DIFFERENTIALLY EXPRESSED DURING DEVELOPMENT.
CC SMALL PROTEIN SPECIES PREDOMINATE IN EARLY EMBRYOS AND ARE
CC UPREGULATED IN COMMITTED MYOBLASTS AND CHONDROCYTES, BUT
CC DOWN-REGULATED UPON TERMINAL DIFFERENTIATION. LARGE SPECIES ARE
CC DETECTED MAINLY IN ADULT TISSUES AND TERMINALLY DIFFERENTIATED
CC CELLS.
CC -!- DOMAIN: ASN AT POSITION 47 OF THE HOMEBOX MAY PARTICIPATE IN
CC REGULATING DNA-BINDING ACTIVITY BY PROMOTING HOMO- AND
CC HETERODIMERIZATION.
CC -!- SIMILARITY: CONTAINS 3 CUT DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEBOX PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X69017; CAA48782.1;
CC PIR; S33121; S33121.
CC HSSP; P10037; LAU7.
CC TRANSFAC; T01485;
CC TRANSFAC; T01985;
CC InterPro; IPR003350; CUT.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF02376; CUT; 3.
CC Pfam; PF00046; homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Transcriptional regulation; Homeobox; DNA-binding;
CC Developmental protein; Nuclear protein; Repeat; Repressor;
CC Coiled coil; Alternative splicing.
CC NON_TER 1 1
CC DNA_BIND <1 67 CUT 1.
CC DOMAIN 113 169 COILED COIL (POTENTIAL).
CC DNA_BIND 374 461 CUT 2.
CC DNA_BIND 557 644 CUT 3.
CC DNA_BIND 684 743 HOMEBOX.
CC SEQUENCE 975 AA; 105428 MW; 724243B32C5BCFD9 CRC64;
Qy 3 VPWELLSALGRSPVLSRLVLPQDATHCS-----PCLSCRLWSDSLCLPGDIVP 55
Db 190 IPSSPMSVSSVYPLALSLLK---PPTAPDTSASTLPNPKKESQDAPGLDLPAAES 246
Qy 56 ACPGVLAPTHQTELVLRCQKETCDCLRVAVHLAVHGHWE-----PE 100
Db 247 AQQ-----VLR-----HVKSGLSGVWKDHWSTVQPERKSAAPPE 283
Qy 101 D--EEKFGAADSG-----V 113
Db 284 DAKSEAGGTKERGGGOGHPIAASSRPHRRSTGRNGPALSPTQSSLSLTGASRS 343
Qy 114 EEPNRLSQ--VWLSFQAYPTARCVLLEVVQVPAALVQFGQSVGVYDCEFAALGSEV 171
Db 344 ETPQNSLPSPSPVPMKPAK-----SVPLTPEQYIYMQEV 383

Query Match 2.8%; Score 104.5; DB 1; Length 975;
Best Local Similarity 20.4%; Pred. No. 2.3;
Matches 174; Conservative 74; Mismatches 244; Indels 363; Gaps 44;

```

```
QY 172 RIWSTQPRYEKELNHTQOLPALPWLNVSDGDNVHLVNLVSEQHFGCLSLYWNQVGP- 230
Db 384 DTIELTRQVREK-----LAKNG-----ICQIFGKVLGLS-----QGV 418
QY 231 -----PKPRWHKNTGPQIITLHTDLVPCLCIQV----- 261
Db 419 SDMLSRKP-WSK-----LTQGREPIRMQLWNLGELGQGVLPVGGQQGQPVLS 468
QY 262 -----PLEPDSVRTNTPFRE-----DRAHQNLWQAARLLTLTQSWLLDAPCSLPARAA 312
Db 469 VTSIQDLQOGCVSSSTPKTSASCSPAPESMSSSESVSLSLT-----ELVQPCP-PIET- 523
QY 313 LWRAPCGDCPCPLVPPLSWENVTVKLEPL-LKGHPNLCVQ--VNSSEKL----- 362
Db 524 -----SKDGKPEPDPDPPAS-----DSQATPPLSGHSALSIOELVAMSPELDITYG 573
QY 363 QLOECLWADSLGP-LKDDVLLETRG-----PODNRSLCALEPSGCTSLPSKAST 411
Db 574 RVKEVLTNNLQORLFGETILGLTQGSVDLLSRPKPHKLSLKGREPF----- 622
QY 412 RAARLGEYLLODQSGQCLOLQDDDLGALMACPMDKYIHKRWALVWLACLLFAAALSIL 471
Db 623 -----VRMQLWLN-----PNN-----VEXLM 639
QY 472 LLKKDHAKAARGAALLLYSADDS-----GFERLVG----- 503
Db 640 DMKRMKAYMKRRHS-----SVDSQCPPEPSVIGIDYSGASPPQHOLKKPRVVLAPEE 695
QY 504 --ALASALCOLPL--RVAVDLWRRELSSAOGPVAFHQAQROTLOEGVVVLLFSPGAV 558
Db 696 KEALKRAYQOKYPSPKTIETELNLTSTVWTFNRYRSIRRE----- 742
QY 559 ALASEWLQDQVSG-PGA-HGPHDAFRASLSVLPDFLQGRAPGSYVGAC--FDRLLHPDA 614
Db 743 -LFEEELQAGSQGAGARHP--SARSS-----GAAPSSGDCDGVAAEGPGA 789
QY 615 VPALFRTPVPTLPSQLPDFLQALQOQPRAPSGRQERAEQVSRALQALDYSFHPPTG- 673
Db 790 ADA-EESAPAAAAKSQ--GGPAEAAVAP-----EEREE-----APRPAEKRRRPRG 835
QY 674 PAPGRGVGPGAGPGA 688
Db 836 PGPRGGGGGPGAPGA 850
RESULT 8
UL06_HSV11 STANDARD; PRT; 676 AA.
AC P10190;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Virion protein UL6.
GN UL6.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dairymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT Herpes simplex virus type 1."
RL J. Gen. Virol. 69:1531-1574(1988).
RN [2]
RP SEQUENCE OF 1-80 FROM N.A.
RX MEDLINE=88091053; PubMed=2826807;
RA McGeoch D.J., Dairymple M.A., Dolan A., McNab D., Perry L.J.,
RA Taylor P., Challberg M.D.;
RT "Structures of herpes simplex virus type 1 genes required for
```

```
RT replication of virus DNA."
J. Virol. 62:444-453(1988).
CC -!- FUNCTION: PRESUMED VIRION PROTEIN; POSSIBLE ROLE IN DNA
CC PACKAGING.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL6,
CC EHv-1 56, EBV BBRF1, HCMV UL104, AND VZV 54.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10879; BAA01652.1; -
DR EMBL; X14112; CAA32342.1; -
DR EMBL; M19119; AAA45818.2; -
DR PIR; F28133; WMBEX6.
DR InterPro; IPR002660; Herpes_UL6.
DR Pfam; PF01763; Herpes_UL6.1.
DR ProDom; PD003210; Herpes_UL6; 1.
SQ SEQUENCE 676 AA; 74091 MW; 8CF9D73C6313FAE9 CRC64;
Query Match 2.8%; Score 103.5; DB 1; Length 676;
Best Local Similarity 22.5%; Pred. No. 1.7;
Matches 102; Conservative 46; Mismatches 153; Indels 153; Gaps 24;
QY 295 TLOSLLDAPCSLPAEAAACWRAPGGD-----PCOPLVPPL--SMENVTVDK 339
Db 130 TLETTLLDFAHGLVA---CF-APGGSPSSPPKYIDMTCTCLGLVPLILKKRQEGGVTOG 184
QY 340 VLEFPLIKGHPNLCVQVNSSEKYLQLOECLWADSLGPLKDDVLLLETRGPQDNRSCLALEP 399
Db 185 LRAF--LKQHP-----LTRQLATVAEA-----AERAGPGFFELALAFD- 220
QY 400 SGCTSLPSKASTRAA-----RLGEYLLQDL---OSGQCLOLWDDDLGALM--- 441
Db 221 -----STRVADYDRVYIYVNHRRGDWLVRDPTISGQGECLVLP-----PLMTGDR 266
QY 442 -----AC-PMDKYIHKRWALVWLACLLFAAALSILLKKDHAKAARGR 485
Db 267 LVFDSPPVQRLFPETVACHSLREHAH-----VCLRLNTASVKVLGKSDSERGVGA 318
QY 486 AALL--LYSADDSGFERLVGALASALCQLPLRVAVDLWLSRRELSAOGPVAFHQAQROT 543
Db 319 ARVNVKVLGEDD---ETKAGSAASRL---VRLINMKGRH-----VGINDTVRSYL 365
QY 544 QEGGVVVLLFSPGAVALCSEWLQDVGSGPGAHGPHDAFRASLSVLPDFLQGRAPGSYVG 603
Db 366 DEAG-----GHL-----IDAPVDG-----TLPGF--GKGNSRGS 394
QY 604 ACDFRLHPDAVPALFRTVPVFTLPSQLPDFL-----GALQOQPRAPSG---RLQERAEQV 656
Db 395 AGODGGRAPQLRQAFRTAVVNNINGVLEGIYNNLFTIERLRETNAGLATQLQERDEL 454
QY 657 SRALQPALDYSFHPPTGPAPGRGVGPGAGPGA 690
Db 455 RRATAGALERQQAADLAESVTGGCGSRPAGAD 488
RESULT 9
W70T_MOUSE STANDARD; PRT; 922 AA.
ID W70T_MOUSE
AC Q9D2V7;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 70 kDa WD-repeat tumor rejection antigen homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

NCBI_TaxID=10090;
 [1]
 QX SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Kidney;
 RC MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aikawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Cariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 FT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -!- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
 CC -!- SIMILARITY: BELONGS TO THE CORONIN FAMILY OF WD-REPEAT PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AK018739; BAB31380.1;
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 6.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR ProDom; PD000018; WD40; 2.
 DR SMART; SM00320; WD40; 6.
 DR PROSITE; PS00678; WD_REPEATS_1; 1.
 DR PROSITE; PS0082; WD_REPEATS_2; 4.
 DR PROSITE; PS0294; WD_REPEATS_REGION; 2.
 KW Repeat; WD repeat.
 FT REPEAT 75 115
 FT REPEAT 124 163
 FT REPEAT 166 205
 FT REPEAT 209 253
 FT REPEAT 539 581
 FT REPEAT 589 629
 FT REPEAT 632 671
 FT REPEAT 725 765
 FT REPEAT 75 765
 SQ SEQUENCE 922 AA; 100860 MW; C8E4BC1EE9CFE242 CRC64;
 Query Match 2.7%; Score 102; DB 1; Length 922;
 Best Local Similarity 21.1%; Pred. No. 3.4;
 Matches 146; Conservative 65; Mismatches 227; Indels 254; Gaps 38;
 QY 130 QAYPTARCVLL-EVQVPALVQFGSGVGVYDVFCEAALGSGVRVSYT--OPRYEKELN 186
 DB 112 EALPSVGVVGLPPELVPEVLFQFHTVDCVLV----STAGTKVTVMDVAKQPLTELEA- 166
 QY 187 HTQQLPALPWLNVSGADGNVHLVL-----NVSEOHFGLSLWY--- 224
 DB 167 HKDLVQSAVW---SRDGAIVGTACKDKQLRIFDPRTAQASQSTQAHENNRDLRLAWTGI 223
 QY 225 -----NQVQGPPEPRWKNLTGPOI--ITLNT--DLVPCV----- 256
 DB 224 QEHLVSTGFNMREARKLWDLTRFSSALASVTLTSPGPLPLDPSGLLVLAGKGEN 283

QY 257 ---CIQVWPLEPDSVTRNICPFRE-----DPRAHQNLMAARLRLTLQSWLL----- 301
 DB 284 QLYCYEVTPOQPALSPVPTQILENVLRGAALVPRALAVMSCEVLQVLQSLDIAPIISH 343
 QY 302 -----DAPCSLPAEAALCWAPGDPQCPPLVPPLSWENVTVDKLVLEFFLL 346
 DB 344 HVPKRAVEFHEDLFDFTAGSMPSADAHMWA--GD-----NQOVKVSLSNPAR 389
 QY 347 KGHPNLCVQVNSSEKLQQLQECLEWAD---SLGPLKDDVLLLE-----TRGPQDNRSIC 395
 DB 390 RHPH-----CFTSSLVPTMEPADWQVPAEMPRADTDLSEGFSSPSL- 432
 QY 396 ALEPSGCTSL-PSKASTR--AARIGEYLLDQSGQCQLQWLDDDLGALWACPMKYIHKR 452
 DB 433 -MSPSTSSLPGLSPSSIGTSPSQRSQSLGLPSC-----KFRHTQ 474
 QY 453 WALVWLACLLFAAALSILLILKKHAKAAAGRAALLYSADDSGFERLVGALASALCOL 512
 DB 475 GS-----LLHRDSHTNLKG--LNLTPGESDGF-----CAN 504
 QY 513 PLRVAVDLWSRRELSSAOGPVAVFAHOR-----ROTLOEGGVVLL-----FSPGAVAL 560
 DB 505 RLRAVAVPL-----LSSGGEVAVLELQKPLRPLDTALPLQNGTAVMDLVDPDPHRLAV 559
 QY 561 CSEWLODVGSGAGHPHDAFRASLSCVLPDLQG--RAPGSYVGCACFDRL-----LHPDA 614
 DB 560 AGE-----DA-RIRLWRVPPGGLNVLTPETVLTGHTKIYLSRFHPLA 603
 QY 615 VPALFR-----TVPVFTLPSPQLPFLGALQOPRAPRSRGLQERAEQV--SRALQP----- 662
 DB 604 ADVLASSYDLTIVIMDL-----QTGAERL-KLQGHQDQIFSLAWSDPGKOLA 650
 QY 663 --ALDSY---FHPPGTAP-GRGVGPGAGPCA 688
 DB 651 TVCKDGHVRVYEPSSPLPLQEGGPGEGRGA 682
 RESULT 10
 ATC7_YEAST
 ID ATC7_YEAST STANDARD; PRT; 1151 AA.
 AC P40527;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Potential phospholipid-transporting ATPase 4 (EC 3.6.3.1).
 GN NE01 OR Y1L048W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Barrrell B.G., Badcock K., Bankier A.T., Bowmen S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RL -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
 CC OF ATP COUPLED WITH THE TRANSPORT OF PHOSPHOLIPIDS (POTENTIAL).
 CC LEADS TO NEOMYCIN-RESISTANCE WHEN OVEREXPRESSED.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IV.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL: Z38060; CAA86174.1; -
DR SGD; S48431; S48431.
DR SGO; S0001310; NEOL.
DR InterPro: IPR001757; E1-E2_ATPase.
DR Pfam: PF00122; E1-E2_ATPase; 1.
DR Pfam: PF00702; Hydrolase; 1.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
FT TRANSMEM 368 388 POTENTIAL.
FT TRANSMEM 417 437 POTENTIAL.
FT TRANSMEM 439 459 POTENTIAL.
FT TRANSMEM 948 968 POTENTIAL.
FT TRANSMEM 971 991 POTENTIAL.
FT TRANSMEM 1021 1041 POTENTIAL.
FT TRANSMEM 1053 1073 POTENTIAL.
FT TRANSMEM 1079 1099 POTENTIAL.
FT TRANSMEM 1110 1130 POTENTIAL.
FT MOD_RES 503 503 PHOSPHORYLATION (PROBABLE).
SQ SEQUENCE 1151 AA; 130217 MW; DC7225CC9577DBE6 CRC64;

Query Match 2.7%; Score 101.5; DB 1; Length 1151;
Best Local Similarity 21.1%; Pred. No. 4.9;
Matches 64; Conservative 49; Mismatches 112; Indels 79; Gaps 15;
QY 102 EEKFGGAADSVETPRNASIAQVYLSQAYPTARCVLLEQVPAALVQFGSGSVYVD 161
Db EDKIQKQVKKSSIELLNAGIRI-W-MLTGDKVETARC-----VSISAKLISRGQYVHTIT-- 801
QY 162 CFEEALGSEVRIWSTYPTRYEKELNHTOOLPALPLNLSA-----DGDNVHLVNLNVEEOH 217
Db 802 -----KVTRP--EGAFNQLYK-----INRNACLLIDGESLGMFLKHVEQEF 842
QY 218 FGLSLYNQV---GGPKPRWH-----KNLGPQIITL-----NHTDLVPLCLIQVW---- 261
Db 843 FDVVVHLPTVIACRTPQKADVALVIRKMTCKRVCICDGDNDYSMIQCADVGVGIVCK 902
QY 262 -----PLEPDSVRNICPFREDPRAHNLW-----QAARLRLTLQSWLLDAPCSLPA 309
Db 903 EGKQASLAADFSIQFCHLE-----LLLHGRNSYKRSKAKLAQFVMHRLIIAICQ--A 955
QY 310 EAALC-----WRAPGDCQPLVPPLSWENVTVDKYLEPPLKKGHPNLCVQVNS 358
Db 956 VYSCSLFEPALYQGLWLVGATCYTMAPVFS-----LTLDHDIIESLTKIYPELYKELTE 1012
QY 359 SEKL 362
Db 1013 GKSL 1016

RESULT 11
UL06_HSVB
ID UL06_HSVB STANDARD; PRT; 753 AA.
AC P28944;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Virion gene 56 protein.
GN 56.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;

RT *The DNA sequence of equine herpesvirus-1.*;
RL Virology 189:304-316(1992).
CC !- FUNCTION: PRESUMED VIRION PROTEIN; POSSIBLE ROLE IN DNA
CC PACKAGING.
CC !- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL6,
CC EHV-1 56, EBV BBRF1, HCMV UL104, AND VZV 54.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M86664; AAB02491.1; -
DR PIR: B36801; WZBER8.
DR InterPro: IPR002660; Herpes_UL6.
DR Pfam: PF01763; Herpes_UL6; 1.
DR ProDom: PD003210; Herpes_UL6; 1.
SQ SEQUENCE 753 AA; 83992 MW; C5E118F78BED203 CRC64;

Query Match 2.7%; Score 100.5; DB 1; Length 753;
Best Local Similarity 22.9%; Pred. No. 3.4; Indels 97; Gaps 20;
Matches 99; Conservative 46; Mismatches 190;
QY 295 TLQSWLLDAPCSLPAEAAACWRAPGDCQPLVPPLSWENVTVDKYLEPPLKKGHPNLCV 354
Db 156 TLQSWLLDFVRSITA---CFSASEPDTGTFASYAKYIDW--IACGLIPLQLKRAPGATV 209
QY 355 QVNSSEKLQEQCLMADSLGPKDDVLLLETRGPQDNRL-----CALE 398
Db 210 HPK-----LWR-----KLPTDVPSELES--CVDERDLAGKLYVANSLLREGLEAVVE 253
QY 399 PSCT-SLPSKASTRAARLGEYLLQDL-----QSGCCLQWD---DDLGLMWACPM 446
Db 254 LARCTASVIMDYDRVNFVHYTRREVAIDSTGKRGECLVLMQPIWKDGSVLFDSPLQ 313
QY 447 KY-----IHKRWALVWLACLLFAAALSILLKDKHAKAAARGRAALLLYSADDSGER 500
Db 314 RICEVCNCHALREHAKLCOLLNTVPVKILVGRKKDEAOGP--GWASKAVDKLWGEEL 371
QY 501 LVGALASALCOLPLRVAVDLWMSRRELSAQGPVAFWFAQRROTLOEGGVVLLFSPGVAL 560
Db 372 HSSSAASRL-----VKLIVNKMSRHIGDITETVRSYLNSTNLISGAQVDTSLPG--- 423
QY 561 CSEWLQDGVSGCAHGP-HDAFRASLSCVLPDFLOGRPGSYVCACFDRLHLPDPAVLF 619
Db 424 ---FGSGKTKQGNMVPQEAFTSVINGMLEG-----YVN-----NLF 462
QY 620 RTVPVFTLPSQLPDFLCAQQAPRSGRLQERAEQVSRALQPALDSYFHPGTPA--PG 677
Db 463 KTIEDLRTGNS-----GLLDQLR-DRESEITHLREQLLRVSOAAADSGTOPGASSAALPG 516
QY 678 RGVGPGAGPGAG 689
Db 517 SGAKSGAG-GLG 527

RESULT 12
CHRD_HUMAN
ID CHRD_HUMAN STANDARD; PRT; 955 AA.
AC Q9H2X0; Q9P022; Q9P023; Q9P024; Q9P025; Q9H2W8; Q9H2W9; Q9H2D3;
AC Q95254;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chordin precursor.
GN CHRD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4).
 RP MEDLINE=21366001; PubMed=11472837;
 RX Millet C., Lemaire P., Orsetti B., Guglielmi P., Francois V.;
 RA "The human chordin gene encodes several differentially expressed
 RT spliced variants with distinct BMP opposing activities.";
 RL Mech. Dev. 106:85-96(2001).
 RN [2]
 RN SEQUENCE OF 115-955 FROM N.A. (ISOFORM 5).
 RP MEDLINE=99000848; PubMed=9782094;
 RX Papanio W.N.G. Scott I.C., Clark T.G., Eddy R.L., Shows T.B.,
 RA Greenspan D.S.;
 RT "Coding sequence and expression patterns of mouse chordin and mapping
 of the cognate mouse chrd and human CHRD genes.";
 RL Genomics 52:236-239(1998).
 RN [3]
 RN SEQUENCE OF 51-125; 705-762; 784-850 AND 872-932 FROM N.A.
 RA Lu B., Bachiller D., Agius E., Larrain J., Piccolo S., Nieters A.,
 RA De Robertis E.M.;
 RT "Bmp binding modules in the chordin patterning protein.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Dorsalizing factor. Key developmental protein that
 CC dorsalizes early vertebrate embryonic tissues by binding to
 CC ventralizing TGF-beta family bone morphogenetic proteins (BMPs)
 CC and sequestering them in latent complexes (by similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- ALTERNATIVE PRODUCTS: 5 isoforms: 1 (shown here), 2, 3, 4 and 5;
 CC may be produced by alternative splicing.
 CC -!- TISSUE SPECIFICITY: Expressed at the highest level in liver.
 CC -!- SIMILARITY: BELONGS TO THE CHORDIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 4 WFEC DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF209928; AAC35767.1; -
 DR EMBL; AF209929; AAC35768.1; -
 DR EMBL; AF209930; AAC35769.1; -
 DR EMBL; AF076612; AAC69835.1; -
 DR EMBL; AF283325; AAC35784.1; -
 DR EMBL; AF136632; AAF70236.1; -
 DR EMBL; AF136633; AAF70237.1; -
 DR EMBL; AF136634; AAF70238.1; -
 DR EMBL; AF136635; AAF70239.1; -
 DR MIM; 603475; -
 DR InterPro: IPR001007; WFEC.
 DR Pfam: PF00093; vwc; 4.
 DR SMART; SM00214; WVC; 4.
 DR PROSITE; PS01208; WFEC; 1.
 DR Developmental protein; Repeat; Glycoprotein; Signal;
 KW Alternative splicing
 FT SIGNAL 1 26
 FT CHAIN 27 955
 FT DOMAIN 51 125
 FT DOMAIN 705 762
 FT DOMAIN 784 850
 FT DOMAIN 872 932
 FT CARBOHYD 217 217
 FT CARBOHYD 351 351
 FT CARBOHYD 365 365
 FT CARBOHYD 434 434
 FT VARSPLIC 85 86
 FT VARSPLIC 87 955
 FT VARSPLIC 85 94
 FT VARSPLIC 95 955
 FT VARSPLIC 328 350

FT VARSPLIC 351 955 MISSING (IN ISOFORM 4).
 FT VARSPLIC 441 480 MISSING (IN ISOFORM 5).
 FT CONFLICT 115 118 RQLP -> QVAA (IN REF. 2).
 FT CONFLICT 189 189 V -> A (IN REF. 2).
 FT CONFLICT 216 216 S -> P (IN REF. 2).
 FT CONFLICT 674 674 T -> P (IN REF. 2).
 SQ SEQUENCE 955 AA; 102013 MW; 12AC030CEACFF3ED CRC64;
 Query Match 2.7%; Score 99; DB 1; Length 955;
 Best Local Similarity 21.2%; Pred No. 6.1;
 Matches 178; Conservative 72; Mismatches 281; Indels 310; Gaps 45;
 QY 2 PVPWFLLSLAL-----GRSPVLSLERLVGP---QDATHCSPLGSC----RLMDSQ 45
 DB 8 PAPLLLLGLLLGSRPARGAGPEPVLPIRSEKEPLVVRGAAGCTFGGKVVYALDETWHPD 67
 QY 46 -----ILCL-----PGDI-----VPAPGPVLAHPLHQLTFLVLR 73
 DB 68 LGPEFGVMRCVLCACEAPQWGRTRTGRVSKNKPCEPTPACG--QPRLPGHCCQT 124
 QY 74 COKETDCDLCRLRAVLAHVG-HWEEPEDEP--KFGGAADSGVEEPRNASLQAQVLSF- 129
 DB 125 CPQE-----RSSERQPSGLSEYPRDPEHRSYDRGEPGAERARGDGHDTFVALLT 177
 QY 130 ---QAYPTARCVLLEQVPAAL-----VQFGOSGVSVVYDCFEAAL--GSEVRI 173
 DB 178 GPRSQAVARARVLLRSSLRFSISYRRLDRTRIRFSDNSGSLVFEPHAPPTQGLVCGV 237
 QY 174 WSTQPYKELNHTQQL-----PALPMLNVASADGNVHLVNVSEEQH 217
 DB 238 WR-APRLSLRLRAEQHLVALVTLTHPSGEVWGVLIRHRAALAAETFSAILTL----- 289
 QY 218 FGLSLYNQVQGPKPWHKNTLGTQITLNT-DLYPCLCIQVWPLEPDSVRTNICPFR 276
 DB 290 -----EGPPQ-----QGVGGITLLTSDTEDSLHLLFLFRLEPRSGGLTQVPLR 335
 QY 277 EDPRAHNLQWAAARLLTLQSWLLDAPCSLPAAALCWAPGDPQCPVPLPSWENVT 336
 DB 336 -----LQILHQGLLREL-----QANVSAQEPGFAEVLV-----NLT 367
 QY 337 VDKV-----LEPFL-LKGHPNLCVQVN-----SSEKLOQECLEWADSLGLPKDD----- 379
 DB 368 VQEMDWLVGELQWALEWAGRPGLRISGHIAARKSCDVLSVLC-GADALIPVQTGAAGS 426
 QY 380 -----VLLLETRGPQ--DNRS-LC---ALEPSGCTSLPSK 408
 DB 427 ASLTLLNGSLIYQVQVVGTSSEYVAMTLETK-PQRREDQRTVLCMHAGLQPGHTAV--- 482
 QY 409 ASTRAARLGEVLLQDLSGOCQLWDDDDLGALWACPMDKYIHKRWALVWLACLLFAAALS 468
 DB 483 -----GIC-----PGLGA-----RGAHMLQNELF----- 502
 QY 469 LIILLKKHAKAARGRAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRRLSA 528
 DB 503 -LNVGKDFPDGELRGHVAALPYCGHSARHDTLPVPLAGALVLPVK-----SQ 550
 QY 529 QGPVAMF-----HAORRQTL-----QEGGVVLLSPGVALCSEWLQGVSGPGAHG 576
 DB 551 AAGHAWLSLTDHCHLHYEVLLAGLGSSEQGTVTAHLGP-----PGTPG 594
 QY 577 PHDAFRASLSCVLPDFLQGRAPGSYVACPDRLHHPDAVPALFR-----TVPVFTLPS 629
 DB 595 PRR-----LLKGFYGSEAQG-----VVKDLEPELLRHLAKGNASLLITTKGS 636
 QY 630 QLPDFLGAQQPRAPRSG--RLOERABQVSRAL-OPALDSVFHP--PGTP--APRGVGP 682
 DB 637 PRGELGQVHIANQCEVGGRLLEAAGAGVRLCAPDTASAAPPVFGLPALAPAKPGGP 696
 QY 683 G 683
 DB 697 G 697

```
RESULT 13
M3K6_HUMAN
ID M3K6_HUMAN STANDARD; PRT; 1011 AA.
AC O95382;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase 6 (EC 2.7.1.1.-).
GN MAP3K6 OR MAPKKK6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99092374; PubMed=9875215;
RA Wang X.S., Diener K., Tan T.-H., Yao Z.;
RT "MAPKKK6, a novel mitogen-activated protein kinase kinase, that
RT associates with MAPKKK3";
RL Biochem. Biophys. Res. Commun. 253:33-37(1998).
CC -!- FUNCTION: ACTIVATES THE JNK, BUT NOT ERK OR P38 KINASE PATHWAYS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF1003118; AAD05304.1;
DR MM; 604468;
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 371 629 PROTEIN KINASE.
FT NP_BIND 377 384 ATP (BY SIMILARITY).
FT BINDING 400 400 ATP (BY SIMILARITY).
FT ACT_SITE 494 494 BY SIMILARITY.
FT ACT_SITE 494 494 BY SIMILARITY.
SQ SEQUENCE 1011 AA; 112512 MW; F7294BA0587D2EC9 CRC64;

Query Match 2.7%; Score 99; DB 1; Length 1011;
Best Local Similarity 20.4%; Pred No. 6.6;
Matches 146; Conservative 93; Mismatches 258; Indels 220; Gaps 39;

QY 113 VEEPRNASLQAQVLSFQAYPTARCVLVQVPAALVQFGSGVYVDFEALGSEVR 172
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 31 VAEQHNVCVHFVTFALNRRNRPGDRAKALSVLLP--LVQLEGSVAPDLY-CMCGRIYKDF 87
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 173 IWSYTOPRYEKLNTQQLPALPWLNVSDGD-NVH-----LVLVNVEOHF----- 218
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 88 FSSGFG-----DAGHREQ--AYHWYRKAFDVEPSLHSGINAAYVLLIAAGQHFEDSKELRL 140
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 219 -GLSL-----YNNQVQGPVKPRWHKNTLGTPOIITLNTHTDLVPC-----L 256
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 141 IGKLGCLLARKGCVEKQWYMDV-----GEYLGAQILANDPTQVLLAAEQLYKL 190
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 257 CIQVWPL-----EPDSVTRNICPFREDP-----RAHONLW-----QAARL 291
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 191 NAPIWLVSVNMTFLYQHFRTPPEPGGPPRAH--FWLHFLQSQCFKTCACAGDQC 248
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 292 RLLTLQSWLLDAPCSLPAEALCWRAPGDCQP-----LVP-----PLSWENVTVDKVL 341
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```



```
FT TRANSMEM 13 33 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 326 346 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 407 427 POTENTIAL.
FT TRANSMEM 445 465 POTENTIAL.
FT TRANSMEM 477 497 POTENTIAL.
FT TRANSMEM 506 526 POTENTIAL.
FT TRANSMEM 535 555 POTENTIAL.
SQ SEQUENCE 587 AA; 63531 MW; F9B74F921BDC8712 CRC64;

Query Match 2.7%; Score 98.5; DB 1; Length 587;
Best Local Similarity 20.0%; Pred. No. 3.6;
Matches 88; Conservative 44; Mismatches 163; Indels 145; Gaps 19;

QY 312 ALCWRAPGDCQPLVPPLSWENVVDKLVLEFPFLKGNLVCQVNS---SEKLQOECL 368
DB 2 ATCW-----PALWAYREFYLIIVLCPLIFLLPLPLIVQTKRAYCAYSIIIMALL 48
QY 369 WADSLGLKDDVL---LLETGPODNRSLCALEPSCCTSLPSKASTRAARLGEYLLQDL 424
DB 49 WCTEALPLAVTVLFPVIVFLLMGIMD-----ASEGLHRLVLSRTPT-----YCLSGA 94
QY 425 QSGCQLQWLDDDLGALWACPMKDIHK-RWALVWLACLLFAAALSILIL----- 471
DB 95 DGGHCRTLEP---AQTHCPPSASYHRSAAALLLGLFMLVTAFLSWISWTATAMVPT 151
QY 472 -----LLKDHAKAA----- 481
DB 152 GHAVLEQLQAKKDVGGNNPTFELQECPOKEVTKLNDGQVPVSAPEPTKTQEHHR 211
QY 482 -ARGRAALLIYSADSGPRLVIGALASALCOLPLRVAVDLWSRRLSQAQPVAFHQR 540
DB 212 FSQGLSLICYSASIGGIDNLGTTPNLVLOGOVN---SLFPKMGCELCFVVMFRLPHH 268
QY 541 QTLQGGVVVLFSPGAVALC---SEWL-----ODGVSPGGAHGPHDAF 581
DB 269 DHLAAG---LAMATGPLPGCOLPELWLWGRGTGKACFFGQHPVQAAGA---HEFC 322
QY 582 RASLSCVLPDFLQ---GRAPSYVG---ACF---DRLLHPDAVPALFRTVPVFTLP 630
DB 323 REDLFTVLVLVLFVLTREPGFFGOWGTVFANEGQSMPSDGTVAIFISLVMFIIPS 382
QY 631 LPDFLALQOQPAPRSGRLQ 650
DB 383 IP---GLMEDPKP--GKLIK 397

RESULT 15
TLEA_RAT
ID TLEA_RAT STANDARD; PRT; 741 AA.
AC Q07141;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Transducin-like enhancer protein 4 (ESP2 protein).
GN TLE4 OR ESP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=94064640; PubMed=8245004;
RA Schmidt C.J., Sladek T.E.;
FT "A rat homolog of the Drosophila enhancer of split (groucho) locus
FT lacking WD-40 repeats.";
RL J. Biol. Chem. 268:25681-25686(1993).
CC -!- FUNCTION: NUCLEAR EFFECTOR MOLECULE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
```

```
CC -!- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L14463; AAC37640.1;
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 5.
CC PRINTS; PR00320; GPROTEINBRPT.
CC SMART; SM00320; WD40; 6.
CC PROSITE; PS00678; WD_REPEATS_1; 2.
CC PROSITE; PS00682; WD_REPEATS_2; 2.
CC PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Nuclear protein; Phosphorylation; Repeat; WD repeat.
FT DOMAIN 1 106 GLN-RICH (OR Q DOMAIN).
FT DOMAIN 107 173 GLY/PRO-RICH (OR GP DOMAIN).
FT DOMAIN 175 242 CCN DOMAIN.
FT REPEAT 441 472 WD 1.
FT REPEAT 499 529 WD 2.
FT REPEAT 543 573 WD 3.
FT REPEAT 585 615 WD 4.
FT REPEAT 667 697 WD 5.
FT REPEAT 708 738 WD 6.
FT MOD_RES 218 218 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD_RES 233 233 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
SQ SEQUENCE 741 AA; 80490 MW; 43F105E568FD87A5 CRC64;

Query Match 2.7%; Score 98.5; DB 1; Length 741;
Best Local Similarity 20.0%; Pred. No. 4.8;
Matches 101; Conservative 59; Mismatches 153; Indels 193; Gaps 26;

QY 23 RLVPQDATHCSGSLSCRLWDSILCLPGDIVPAG---PVLAP---THLQTELVLRCQKET 78
DB 297 RTDAPTGSNSTPGLR-----PVPKPPGVDPLASSLRTMAVPCPYPT 340
QY 79 DCDLCLVAVHLAVGHWEPEDEKEFGGAADSGVEEPRNASLQAAQVLSFQAYPTARC 138
DB 341 P-----FGIVPHAGMNGELTSP-----GAAYAGL---HNISPMQSA--- 376
QY 139 LLEVQVPAALVQFGQSVGVVYDCFEAALGSEVRISWYTPQRYEKLNHTQQLPALP--- 195
DB 377 -----AAAAAAAYGRS-----PVGFDPHHMRVPAIPNL 407
QY 196 -----WLVNSADGDNVHLVNLVNSEQHGLSLVWNVQVGPVPRHKNLTGP--- 242
DB 408 TGIPGKPAYSFHVSADG-----OMQPVPEP---PDALITGP 443
QY 243 -----QIITLHNTDLVPCPL-----CLOVWPLEPDSVVRNICFFREDPRAHQ 283
DB 444 RHARQINTLHNGEVVCAVTSINPRHRYVTGGKGVKVVWDITDPGNKSPVSOL---DCLNRD 501
QY 284 NLWQAARL---RLL-----TLQSWLLDAPCSLPAEALCWAPGDCQPLVPLSW 332
DB 502 NYIRSCRLLPDGRLVIGGEASTLSINDLAAP-TPRIKAELTSSAPA---CYALA----- 552
QY 333 ENVTVDKLVLEFPFLKGNLVCQVNSSEKILQLOECLWADSLGPKLDDVLLLETGPODNR 392
DB 553 --ISPDSKVCVF-----SCCSDGNIA-----VW-----DLHNOITLVRFQGHDTGA 590
QY 393 SLCALEPSCCTSLPSKASTRAARLGEYLLQDLQSGCQLQWLWD---DDLGAWACPMKDIH 450
DB 591 SCIDISNDG-TKLMTGGLDNTR-----SWDLREGQLQOHDTFSQIFSLGCTPGE--- 641
QY 451 KRWALVWLACLLFAAALSILILKKD 476
```


Tuế May 21 14:54:59 2002

us-09-608-918-2.rsp

Page 15

Db 642 -----WLAUGMENSNEVLHVTKPD 661

Search completed: May 19, 2002, 17:01:01
Job time: 10242 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 11:19:39 ; Search time 89.34 seconds
(without alignments)
860.343 Million cell updates/sec

Title: US-09-608-918-2

Perfect score: 3716

Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPGRGVGPGAGDGT 692

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID55/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID55/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID55/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID55/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID55/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID55/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID55/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID55/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID55/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID55/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID55/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID55/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID55/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID55/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID55/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID55/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID55/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID55/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID55/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID55/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID55/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3716	100.0	692	22	AA61880 Human cytokine rec
2	3699.5	99.6	705	22	AAU29322 Human PRO polypept
3	3699.5	99.6	705	22	AAU04956 Human Interleukin
4	3699.5	99.6	705	22	AB87606 Human PRO20040..H
5	3699.5	99.6	705	22	AB87606 Human PRO20040..H
6	3604.5	97.0	675	22	AA61884 Chimeric zcytor14
7	3588	96.6	688	22	AA61885 Chimeric zcytor14
8	2971	80.0	575	22	AA61881 Human variant zcyt
9	1432	38.5	309	21	AA76143 Human secreted pro
10	1305	35.1	267	22	AB88448 Human membrane or
11	1300	35.0	332	22	AA25864 Human protein sequ

12	1079	29.0	204	22	AAE10920 Human gene l2 enco
13	326.5	8.8	617	21	AA76048 Murine skin cell p
14	326.5	8.8	617	22	AB55987 Skin cell protein,
15	309.5	8.3	667	22	AAU04957 Human Interleukin
16	174	4.7	866	17	AAU04185 Human Interleukin
17	174	4.7	866	19	AAW61272 Human Interleukin
18	174	4.7	866	20	AAW92409 Human IL-17R prote
19	174	4.7	866	21	AAV99941 Human IL-17R prote
20	174	4.7	866	21	AAV97131 Human Interleukin
21	174	4.7	866	21	AAV97181 Human Interleukin
22	174	4.7	866	21	AAU03807 Human Interleukin
23	174	4.7	866	22	AAU03807 Human IL-17R (hCTL
24	174	4.7	866	22	AAU03807 Human Interleukin
25	165	4.4	864	17	AAW04184 Murine Interleukin
26	165	4.4	864	19	AAW61271 Mouse Interleukin
27	165	4.4	864	20	AAW92408 Murine IL-17R prot
28	165	4.4	864	21	AAV99935 Murine IL-17R prot
29	165	4.4	864	21	AAV97130 Murine Interleukin
30	165	4.4	864	21	AAV97180 Murine Interleukin
31	165	4.4	864	21	AAU03806 Murine Interleukin
32	165	4.4	864	22	AAU03806 Murine IL-17R poly
33	165	4.4	864	22	AAU03806 Murine Interleukin
34	162	4.4	330	21	AAV75947 Murine skin cell s
35	162	4.4	330	22	AB55886 Skin cell protein,
36	119.5	3.2	840	21	AAV93652 A mammalian solubi
37	119.5	3.2	840	21	AAV70028 Soluble Interleuki
38	118	3.2	1042	22	ABG26706 Novel human diagno
39	116.5	3.1	1711	19	AAW70506 Osteotesticular pr
40	116.5	3.1	1711	19	AAW70507 Mutant osteotestic
41	114.5	3.1	336	22	ABG03482 Novel human diagno
42	112.5	3.0	1208	20	AAW95050 Human helicase pro
43	112.5	3.0	1208	21	AAU20993 Human RecQ4 helica
44	110	3.0	199	22	ABG20476 Novel human diagno
45	110	3.0	664	22	AAE06034 Mouse oocyte prote

ALIGNMENTS

RESULT 1

AA61880

ID AA61880 standard; Protein; 692 AA.

XX AC AA61880;

XX DT 08-MAY-2001 (first entry)

XX DE Human cytokine receptor Zcytor14.

XX KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;

XX KW antiinflammatory; gene therapy; vaccine.

XX OS Homo sapiens.

XX PN WO200104304-A1.

XX PD 18-JAN-2001.

XX PF 30-JUN-2000; 2000WO-US18383.

XX PR 07-JUL-1999; 99US-0348854.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Presnell SR, Burkhead SK, Pownder SL;

XX XX WPI; 2001-112618/12.

XX DR N-PSDB; AAC85027.

XX XX New polypeptide encoding a human cytokine receptor Zcytor14, for treating inflammation e.g. rheumatoid arthritis.
XX Claim 2; Page 2; 112pp; English.

XX CC The invention provides a new human cytokine receptor designated zcytor14.
CC zcytor14 can be expressed by standard recombinant methodology. The
CC encoding nucleic acid is useful for detecting the expression of a
CC zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be
CC used to screen biological samples in vitro for the presence of zcytor14.
CC Proteins, polypeptides and peptides having zcytor14 activity can be
CC administered to a subject who lacks an adequate amount of this
CC polypeptide, for treating inflammation and conditions such as rheumatoid
CC arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14
CC antibodies) can be used to treat a subject who produces an excess of
CC zcytor14. zcytor14 nucleotide sequences can also be used to provide
CC zcytor14 to a subject. The present sequence represents the human
CC cytokine receptor zcytor14.
XX
SQ Sequence 692 AA;

Query Match 100.0%; Score 3716; DB 22; Length 692;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVPWLLSLALGRSPVLSRLVGPQDATHCSPLGSLRWLSDILCLPGDIVPAGPV 60
DB 1 MPVPWLLSLALGRSPVLSRLVGPQDATHCSPLGSLRWLSDILCLPGDIVPAGPV 60
QY 61 LAPTHLQTELVLRCQETDCDLRLVAVHLAVHGHWEPEDEEFKFGAADSVEEPNRS 120
DB 61 LAPTHLQTELVLRCQETDCDLRLVAVHLAVHGHWEPEDEEFKFGAADSVEEPNRS 120
QY 121 LQAQVVLSTQAYTARCVLEVVQVPAALVQFGSVGVYDCCFAALGSEVRIWSYTPR 180
DB 121 LQAQVVLSTQAYTARCVLEVVQVPAALVQFGSVGVYDCCFAALGSEVRIWSYTPR 180
QY 181 YKELNHTQQLPALPWLNSADGNVHLVNVSEQHFGLSLYNQVQGGPKPRWKNLT 240
DB 181 YKELNHTQQLPALPWLNSADGNVHLVNVSEQHFGLSLYNQVQGGPKPRWKNLT 240
QY 241 GPQITLNTDVLPCICVWPLEPDSVRTNICPFREDPRAHONLWQAAARLLTLQSWL 300
DB 241 GPQITLNTDVLPCICVWPLEPDSVRTNICPFREDPRAHONLWQAAARLLTLQSWL 300
QY 301 LDAPCSLPAEALCWAPGDDPCQLVPPLSWENVTVDKVFLLKGHPLNLCVQVNSSE 360
DB 301 LDAPCSLPAEALCWAPGDDPCQLVPPLSWENVTVDKVFLLKGHPLNLCVQVNSSE 360
QY 361 KLQQLCLWADSLGPKLDDVLLLETRGPQNRSLCALPESGCTSLPSKASTRAARLGEYL 420
DB 361 KLQQLCLWADSLGPKLDDVLLLETRGPQNRSLCALPESGCTSLPSKASTRAARLGEYL 420
QY 421 LQDLOSQCLQWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALLSLILLKDHAKA 480
DB 421 LQDLOSQCLQWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALLSLILLKDHAKA 480
QY 481 AARGRAALLYLSDSGFERLVGALASALCQLPLRVAVDLWSRRELQAQGPVAFHQAORR 540
DB 481 AARGRAALLYLSDSGFERLVGALASALCQLPLRVAVDLWSRRELQAQGPVAFHQAORR 540
QY 541 QTLQEGGVVVLFPSPGAVALCSEWLQDGVSGPAGHPHDAFRASLSCVLPDFLQGRAPGS 600
DB 541 QTLQEGGVVVLFPSPGAVALCSEWLQDGVSGPAGHPHDAFRASLSCVLPDFLQGRAPGS 600
QY 601 YGACFDRLLHPDAVPALEFVTVFVLPQLPDLFLGALQOPRAPRSRLQERAEQVSRAL 660
DB 601 YGACFDRLLHPDAVPALEFVTVFVLPQLPDLFLGALQOPRAPRSRLQERAEQVSRAL 660
QY 661 QPALSDFRHPPTPAPGRGVGPGAGPCAGDGT 692
DB 661 QPALSDFRHPPTPAPGRGVGPGAGPCAGDGT 692

RESULT 2
AAU29322

ID XX AAU29322 standard; Protein; 705 AA.
AC XX AAU29322;
DT XX 18-DEC-2001 (first entry)
XX XX Human PRO polypeptide sequence #299.
DE PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX OS Homo sapiens.
XX PN WO200168848-A2.
XX PD 20-SEP-2001.
XX PF 28-FEB-2001; 2001WO-US06520.
XX PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194677P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 18-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX XX (GETH) GENENTECH INC.
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX WPI: 2001-602746/68.
XX N-PSDB: AAS46223.
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX

PS Claim 11; Fig 598; 774pp; English.

xx Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.

CC The PRO polypeptides and their associated nucleic acids can be used to

CC detect the presence of a tumour in a mammal by comparing the level of

CC expression of a PRO polypeptide in a test sample of cells from the animal

CC and a control sample of normal cells, whereby a higher level of

CC expression in the test sample indicates the presence of a tumour in the

CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats

CC and rabbits but are preferably human. The polypeptides can be used to

CC stimulate tumour necrosis factor (TNF) alpha release from human blood,

CC when contacted with it. A specific polypeptide can be used to stimulate

CC the proliferation or differentiation of chondrocyte cells. The PRO

CC proteins can be used to determine the presence of tumours and also

CC susceptibility to tumour development, particularly adrenal, lung, colon,

CC breast, prostate, rectal, cervical, or liver tumours, in mammalian

CC subjects. The oligonucleotide probes specific for the PRO nucleic acids

CC can be used for genetic analysis of individuals with genetic disorders.

xx SQ Sequence 705 AA;

Query Match 99.6%; Score 3699.5; DB 22; Length 705;

Best Local Similarity 98.2%; Pred. No. 0;

Matches 692; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 1 MPVPFLLSLALGRSPVLSRLERLVPQDATHCSPGLSCLRLWSDILCLPGDIVPAPGPV 60

Db 1 mpvpfllslalgrspvlsrlervpqpqathcspglscrlwdsdlclpgdivpapgpp 60

QY 61 LAPTHLOTELVLRCOKETDCDCLRVAVLHVHGWEEPEDEKFGGAADSGVEEPRNAS 120

Db 61 lapthlqtelvlrcoketdcdclrvavhlavhgwepedeekfkggaadsgveeprnas 120

QY 121 LQAOVYLSFQAYPTARCYLEVQVPAALVQFGOSGVSVVYDCFEAALGSEVRISWYTPQR 180

Db 121 lqaovylsfqayptarcylevqvpaalvqfgosgvsvvdydcfeaalgsevrirwstqpr 180

QY 181 YEKELNHTQOLPALPMLNVSAGDNVHLVNLVSEQHFGLSLVNVOGQPKPRWHKNTL 240

Db 181 yekehnhtqolpalpmlnvsagdnvhlvlnvseeqhfglslyvnvgqgpkprwhkntl 240

QY 241 GPQITLNTHTDLVPCICIOVPLEPDSVRTNICPFREDPRAHONLWQAARLRLTLQSWL 300

Db 241 gpqitlnthtdlvpcicigvplepdsrvtrnicpfredprahnlwqaarlrltlqswl 300

QY 301 LDAPCSLPAEALCWRAPGDCPCQPLVPPLSWENVTVDKLEPLLKHPNLCVQVNSSE 360

Db 301 ldapcslpaealcwragpdcpcqplvpplswenvtdklevpllkghpnlcvqvnssse 360

QY 361 KLOLQECILWADSLGPKLDDVLLLETRGPQDNRSICALPEPGCTSLPSKASTRAARLGEYL 420

Db 361 kllqleclwadsllgplkddvllletrgpqdnrsicalpepgctslpskastaarlgelyl 420

QY 421 LODLQSGCQLQWDDDLGALWACPMKDYIHKRWALVWLACLLFAAALSILLKLDHAK - 479

Db 421 lodlqsgcqlqlwdddlgalwacpmkdyihkrwalvwlacillfaaalsillkldhakg 480

QY 480 -----AAARGRAALLYSDSGFERLVGALASALCOLPLRVAVDLWSRRRLS 527

Db 481 wlrlilkdvrsrgaaargraallysddsgferlvgalasalcolplrvavdlwsrrrls; 540

QY 528 AQGPVAFHQAQRRTLOEGGVVLLFSPGAVALCSEMLODVGSPGAGHPHDAFRASLSC 587

Db 541 aggpvafhaqrtrtloeggvvllfspgavalcsewlvqgvsgpghgphdafraslsc 600

QY 588 VLPDFLQGRAPSGYVACFDRLLHPDAVPALFRTVPVFTLPSQLPDFLQALQOPRPSG 647

Db 601 vlpdfllqgrapsgyvacfdrllhpdavpalfrtvpvftlpsqlpdfllqalqgrprsg 660

QY 648 RLQERAEQVSRALQALPDLSYFHPHGPAPGRGVGPGAGCGAGDGT 692

Db 661 rlqeraeqvsralqalpdlsyfhpghgpapgrgvpgagpgagdggt 705

RESULT 3

AAU04956

XX ID AAU04956 standard; Protein; 705 AA.

XX AC AAU04956;

XX DT 24-OCT-2001 (first entry)

XX Human Interleukin 17 receptor, IL-17RH2.

XX Human; Interleukin-17 receptor; IL-17RH2; agonist; antagonist;

XX Human; DNA 164625-2890; systemic lupus erythematosus;

XX rheumatoid arthritis; osteoarthritis; diabetes mellitus;

XX allergic disease; asthma; demyelinating disease;

XX degenerative cartilaginous disorder; transplantation associated disease.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein 21..705

FT Region 107..112

FT Modified-site /note= "N-myristoylation site"

FT Region 118..121

FT Modified-site /note= "Asn is N-glycosylated"

FT Region 152..157

FT Modified-site /note= "N-myristoylation site"

FT Modified-site 186..189

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 198..201

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 211..214

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 238..241

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 248..251

FT Modified-site /note= "Asn is N-glycosylated"

FT Region 319..324

FT Modified-site /note= "N-myristoylation site"

FT Modified-site 334..337

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 357..360

FT Modified-site /note= "Asn is N-glycosylated"

FT Region 391..394

FT Modified-site /note= "Asn is N-glycosylated"

FT Region 438..443

FT Domain 453..473

FT Region 516..521

FT Region 552..555

FT /note= "cAMP/GMP-dependent protein kinase phosphorylation site"

FT Region 583..586

FT /note= "Glycosaminoglycan attachment site"

FT Region 612..617

FT Region 692..697

FT Region 696..701

FT Region /note= "N-myristoylation site"

FT Region 700..705

FT /note= "N-myristoylation site"

PN WO200146420-A2.

PD 28-JUN-2001.

XX

DR N-PSDB; AAF92138.
XX Eighty four nucleic acids encoding PRO polypeptides, useful in
PT molecular biology, including use as hybridization probes, and in
PT chromosome and gene mapping.
XX
XX
XX Claim 12; Fig 162; 278pp; English.
XX
XX The present sequence is a human PRO polypeptide (secreted and
CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
CC anti-PRO antibodies are useful for preparation of a medication useful in
CC the treatment of a condition which is responsive to the PRO protein,
CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
CC employed as molecular weight markers for protein electrophoresis. The PRO
CC coding sequence has applications in molecular biology, including use as
CC hybridisation probes, and in chromosome and gene mapping.
XX
XX
SQ Sequence 705 AA;

Query Match 99.6%; Score 3699.5; DB 22; Length 705;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 692; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

Qy 1 MPVPWFLSLALGRSPVLSRLVGPQDATHCSGLSCRLWSDILCLPGDIVPAPGPV 60
Db 1 mpvpwflslalgrspvlsrlvvgpqdathcspslscrlwdsdlclpgdivpapgvp 60

Qy 61 LAPTHLOTELVRCKQKEDCDCLRLVAVHLAVHGHWEPEDEKFGGAADSGVEEPRNAS 120
Db 61 lapthlqtelvircqketcdclrlvavhlavhghwepeedeekfggaadsgveeprnas 120

Qy 121 LOAQVWLSFOAYPTARCVLLEVQVPAALVQFGSGSVVYDCFEALGSEVRWISYTOPR 180
Db 121 lqaqvwlslfoayptarcvllvqvpaalvqfgsgsvvydcfealgsvevriwstqpr 180

Qy 181 YKELNHTQOLPALPWLNVSDGDNVHLVNVSEQHFGLSLYWNOVGPPKPRWHKNTL 240
Db 181 ykelnhhtqolpalpwlrvnsdgdnvhlvlnvseeqhfglsllywnovgppkprwhkntl 240

Qy 241 GPQIITLNHTDLVPCLCIQVWPLPDSVRTNICPFREDPRAHQLNQAAARLLTLTQSWL 300
Db 241 gpqiitlnhtdlvpcliciqvwplpdsvrtnicpfredprahqlnqaarlltltqswl 300

Qy 301 LDAPCSLPAEALCWRAPGDCQPLVPPLSWENVTVDKVLPEPLKLGHPNLCVQVNSSE 360
Db 301 ldapcslpaealcwragpgdcqplvpplswenvtvdkvlefpilkghpnlcqvnsse 360

Qy 361 KIQLOECLWADSLGPLKDDVLLLETRGPQDNRSICALEPSGCTSLPSKASTRAARLGEYL 420
Db 361 kqlqleclwadslgplkddvllletrgpqdnrsicalepsgctslpskasttraarlgelyl 420

Qy 421 LQDLQSGQCQLQWDDDLGALWACPMDBKYIHKRWALVWLACLLFAAALSLTLKDKHAK - 479
Db 421 lqdlsgqcqlqlddldgalwacpmdbyihkrwalvwlacllfaaalsltllkdkhakg 480

Qy 480 -----AAARGRAALLYSADDSGFERLVGALASALCOLPLRVAVDLWSRRELS 527
Db 481 wlrlkqdvrsaaaraallysaddsgferlvgalasalcqlprvavdlwsrrels 540

Qy 528 ACPVAVFHAQRQRTLOEGGVVWLLFSPGAVALCSEWLQDVGSGPGANGPHDAFRASLSC 587
Db 541 adgpvavfhaqrtrtliqeggvvllfsgavalcsewlgdvgsgpgangphdfraslsc 600

Qy 588 VLPDFLQGRAGPSYGCACFDRLHHPDAVPALFRTPVETPLPSOLPDLFGALQOPRAPRSG 647
Db 601 vlpdfqgragsygcacfdrlhpdavpalfrtpvetplpsolpdlfgalqoprprsg 660

Qy 648 RLQERAEQVSRALQALPDLSYFHPPTGPAFGRGVGPAGPGAGDGT 692
Db 661 rlqeraeqvsralqalpdlsyfhpptgpafrgrgvpgagpgagdgdt 705

RESULT 5
AAB61884
ID AAB61884 standard; Protein; 705 AA.
XX
AC AAB61884;
XX
DT 08-MAY-2001 (first entry)
XX
DE Chimeric Zcytor14 protein #2.
XX
KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;
KW antiinflammatory; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN W0200104304-A1.
XX
PD 18-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18383.
XX
PR 07-JUL-1999; 99US-0348854.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Presnell SR, Burkhead SK, Pownder SL;
XX
DR WPI; 2001-112618/12.
XX
PT New polypeptide encoding a human cytokine receptor Zcytor14, for
PT treating inflammation e.g. rheumatoid arthritis -
XX
PS Claim 2; Page 105-107; 112pp; English.
XX
CC The invention provides a new human cytokine receptor designated Zcytor14.
CC Zcytor14 can be expressed by standard recombinant methodology. The
CC encoding nucleic acid is useful for detecting the expression of a
CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be
CC used to screen biological samples in vitro for the presence of Zcytor14.
CC Proteins, polypeptides and peptides having Zcytor14 activity can be
CC administered to a subject who lacks an adequate amount of this
CC polypeptide, for treating inflammation and conditions such as rheumatoid
CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14
CC antibodies) can be used to treat a subject who produces an excess of
CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide
CC Zcytor14 to a subject. The present sequence represents a chimeric
CC Zcytor14 protein.
XX
SQ Sequence 705 AA;

Query Match 99.6%; Score 3699.5; DB 22; Length 705;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 692; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

Qy 1 MPVPWFLSLALGRSPVLSRLVGPQDATHCSGLSCRLWSDILCLPGDIVPAPGPV 60
Db 1 mpvpwflslalgrspvlsrlvvgpqdathcspslscrlwdsdlclpgdivpapgvp 60

Qy 61 LAPTHLOTELVRCKQKEDCDCLRLVAVHLAVHGHWEPEDEKFGGAADSGVEEPRNAS 120
Db 61 lapthlqtelvircqketcdclrlvavhlavhghwepeedeekfggaadsgveeprnas 120

Qy 121 LOAQVWLSFOAYPTARCVLLEVQVPAALVQFGSGSVVYDCFEALGSEVRWISYTOPR 180
Db 121 lqaqvwlslfoayptarcvllvqvpaalvqfgsgsvvydcfealgsvevriwstqpr 180

Qy 181 YKELNHTQOLPALPWLNVSDGDNVHLVNVSEQHFGLSLYWNOVGPPKPRWHKNTL 240
Db 181 ykelnhhtqolpalpwlrvnsdgdnvhlvlnvseeqhfglsllywnovgppkprwhkntl 240

Qy 241 GPQIITLNHTDLVPCLCIQVWPLPDSVRTNICPFREDPRAHQLNQAAARLLTLTQSWL 300
Db 241 gpqiitlnhtdlvpcliciqvwplpdsvrtnicpfredprahqlnqaarlltltqswl 300

Db 241 gpqiitlnhtdlvpcleicvwwplepsdvrtncifredprahqnlwqaarlrltlqswl 300
 QY 301 LDAPCSLPAEALCWRAFGDPCQPLVPPLSWENVTVDKVLFEPLLLKGPMLCQVNSSE 360
 Db 301 ldpacsipaeaalcwragpgdpcqplvpplswenvtvdkvlefpplkghpnlcvqnsse 360
 QY 361 KLQLECLWADSLGPKDDVLLLETRGPQDNRSICALPSGCTSLPSKASTRAARLGEYL 420
 Db 361 klqlqeclwadsigplkddvllletrgpqdnrsicalpsgctsipskastaaralgeyl 420
 QY 421 LODLOSQCQLQWDDDLGALWACPMKDYIHKRWALVWLACLLFAAALSLILLKKDHAK - 479
 Db 421 lqdlqsgqcqlqwdldlgalwacpmdkyihkrwalvwlacllfaaalslillkkdhakg 480
 QY 480 -----AAAGRAALLYSDDSGFERLVGALASALCQPLRVAVDLMSRRRLS 527
 Db 481 wlrlklqdvrtsgaaargaalylsadsdgsferlvgalasalclprrvavdlmsrrrls 540
 QY 528 AOGPVAFHQAORRTLOEGGVVLLFSPGVALCSEWLQDGVSGPGAHGPHDAFRASLSC 587
 Db 541 aggpvafhaqrtrtlqeggvvllfsgpvalcsewlqdgvsppgahghphdafrafrasc 600
 QY 588 VLPDFLQGRAPSGVYGACFDRLLHPDAVPALFRTPVFTLPSQLPDLFGALQOPRPSG 647
 Db 601 vlpdfiqgrapsgvygacfdrlhhpdavpalfrtvpvftlpsqlpdfalqopraprsg 660
 QY 648 RLQERAEQVSRLQPALDSVFHPPGTPAGRGVSGPAGAGDGT 692
 Db 661 rlqeraeqvsralqpaldsyfhhpgtpagrgvgpgagagdggt 705

RESULT 6

AAB61885
 ID AAB61885 standard; Protein; 675 AA.

XX AC AAB61885;

XX DT 08-MAY-2001 (first entry)

XX DE Chimeric Zcytor14 protein #3.

XX KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;
 XX KW antinflammatory; gene therapy; vaccine.

XX OS Homo sapiens.

XX PN WO200104304-A1.

XX PD 18-JAN-2001.

XX PF 30-JUN-2000; 2000WO-US18383.

XX PR 07-JUL-1999; 99US-0348854.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Presnell SR, Burkhead SK, Powder SL;

XX DR WPI; 2001-112618/12.

XX PT New polypeptide encoding a human cytokine receptor Zcytor14, for
 XX PT treating inflammation e.g. rheumatoid arthritis -

XX PS Claim 2; Page 107-109; 112pp; English.

XX CC The invention provides a new human cytokine receptor designated Zcytor14.
 CC Zcytor14 can be expressed by standard recombinant methodology. The
 CC encoding nucleic acid is useful for detecting the expression of a
 CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be
 CC used to screen biological samples in vitro for the presence of Zcytor14.
 CC proteins, polypeptides and peptides having Zcytor14 activity can be
 CC administered to a subject who lacks an adequate amount of this
 CC polypeptide, for treating inflammation and conditions such as rheumatoid

CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14
 CC antibodies) can be used to treat a subject who produces an excess of
 CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide
 CC Zcytor14 to a subject. The present sequence represents a chimeric
 CC Zcytor14 protein.

XX SQ Sequence 675 AA;

Query Match 97.0%; Score 3604.5; DB 22; Length 675;
 Best Local Similarity 97.5%; Pred. No. 0;
 Matches 675; Conservative 0; Mismatches 0; Indels 17; Gaps 1;

QY 1 MPVPWFLLSLALGRSPVWLSRLVGPDPDATHCSFGLSCRLWDSITLCLPGLDVPAGPV 60
 Db 1 mpvpwfllslalgrspvwlslrlvqpdpdathcsfglscrlwdsitlclpgldivpapg 60
 QY 61 LAPTHLOTELVLRCQKEDTCDLCLRVAVHLAVHGHWEPEDEKFGGAADSGVEEPNRS 120
 Db 61 lapthlqtelvlrcqketcldclrvavhlavhghwepeedeekfggaadsgveepnrs 120
 QY 121 LQAQVVLVFOAYPTARCVLLEVOVPAALVQFGQSVGVVYDFEALGSEVRWSTQPR 180
 Db 121 lqaqvvlvfoayptarcvllvovpaalvqfgqsvgvvydfcealgsvevrwstqpr 180
 QY 181 YEKELNHTQOLPALPMLNVSADGDNVHLVNSEQHFLSLYWNQVQPPKPRWHKNT 240
 Db 181 yekelnhtqolpalpmlnvsadgdnvhlvnseeqhfslslywnvqvgppkprwhknt 240
 QY 241 GPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHONLWQAARLLTLQSWL 300
 Db 241 gpqiitlnhtdlvpcleicvwwplepsdvrtncifredprahqnlwqaarlrltlqswl 300
 QY 301 LDAPCSLPAEALCWRAFGDPCQPLVPPLSWENVTVDKVLFEPLLLKGPMLCQVNSSE 360
 Db 301 ldpacsipaeaalcwragpgdpcqplvpplswenvtvd-----vnsse 343
 QY 361 KLQLECLWADSLGPKDDVLLLETRGPQDNRSICALPSGCTSLPSKASTRAARLGEYL 420
 Db 344 klqlqeclwadsigplkddvllletrgpqdnrsicalpsgctsipskastaaralgeyl 403
 QY 421 LODLOSQCQLQWDDDLGALWACPMKDYIHKRWALVWLACLLFAAALSLILLKKDHAKA 480
 Db 404 lqdlqsgqcqlqwdldlgalwacpmdkyihkrwalvwlacllfaaalslillkkdhaka 463
 QY 481 AARGRAALLYSDDSGFERLVGALASALCQPLRVAVDLMSRRRLSACQPVAFHQAORR 540
 Db 464 aargraallysadsdgsferlvgalasalclprrvavdlmsrrrlsacqpvafhagrr 523
 QY 541 QTLQEGGVVWLLFSPGVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDLQGRAPGS 600
 Db 524 qtlqeggvvllfsgpvalcsewlqdgvsppgahghphdafrafrascvlpdlqgrapgs 583
 QY 601 YVGACFDRLLHPDAVPALFRTPVFTLPSQLPDLFGALQOPRPSRGLQERAEQVSRL 660
 Db 584 yvgacfdrlhhpdavpalfrtvpvftlpsqlpdfalqopraprsgrlqeraeqvsral 643
 QY 661 QPALDSYFHPGTPAGRGVSGPAGAGDGT 692
 Db 644 qpaldsyfhhpgtpagrgvgpgagagdggt 675

RESULT 7

AAB61883
 ID AAB61883 standard; Protein; 688 AA.

XX AC AAB61883;

XX DT 08-MAY-2001 (first entry)

XX DE Chimeric Zcytor14 protein #1.

XX KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;

Query Match		80.0%;	Score 2971;	DB 22;	Length 575;
Best Local Similarity		94.9%;	Pred. No. 9.6e-263;		
Matches 562;		Conservative 0;	Mismatches 0;	Indels 30;	Gaps 2;
QY	114	EEPRNASLQAOVVLVSFOAYPTARCVLLEVOVPAALVFGVSGSVVYDCFEAALGSEVRI	173		
Db	1	eeprnaslqavvlsfqaayptarcvilevqvpaaalvqfgsvsvvydcfeaalgservri	60		
QY	174	WSYTPQRYEKELNHTQOLPALPWLNSADGDNVHLVNLNVEEQHFGLSLYWNOVGPPKP	233		
Db	61	wsytqpryekelnhtqolpalpwlinsadgdnvhlvlnvseeqhfqislywnvgvppkp	120		
QY	234	RWHKNLTGPQIITLNTDVLVPCICIQVWPLEPDSVRTNICPFREDPRAHQNLMOAARLRL	293		
Db	121	rwhknltpqiiitlnhtdlvpcicqivwplepsdvrtnicpfredprahqnlwqaarlrl	180		
QY	294	LTLQSWLLDAPCSLPAEALCWAPAGDPCOPLVPPLSWENVTVDKVLFFPLKAGHPNLC	353		
Db	181	ltlqswlldapcslpaeaalcwapagdpocplvpplswenvtvd-----	225		
QY	354	VQVNSSEKLQLECLWADSLGPKLDVLLLETRGPQDNRSILCALEPSGCTSLPSKASTRA	413		
Db	226	--vnseklqlqecelwadsigplkddvllletrgpqdnrsilcalepsgctsipskastro	283		
QY	414	ARLGEYLLQDLQSGQCLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLILL	473		
Db	284	arlgeyllqdlqsgqclqldddlgalwacpmdkyihkrwalvwlacllfaaalslilll	343		
QY	474	KDHAK-----AAAGRAALLYSADDSGFERLVGALASALCOLPLRVAVDL	520		
Db	344	kdhakgwrlrkqdvrsagaarallaysadsgferlvgalasalcqlplravadl	403		
QY	521	WSRRELSAQGPVAMFHAQRQTTLQEGGVVLLFSPGAVALCSEWLQDGVSGPAGHPHDA	580		
Db	404	wsrrelsaqgpvamfhaqrqtllqeggvvllfspgavalcsewldgvsgpaggphda	463		
QY	581	FRASLSCVLPDFLQGRAPGSYVGACFDRLHPDAVPALEFRTVPVFTLPSPQLPDPFLGALQQ	640		
Db	464	fraslscvlpdfllqgrapgsyvgacfdrlhpdavpalftrtpvftlpqlpdpflgalqq	523		
QY	641	PRAPSGRLQERABQVSRALQALDSYFHPGTPAPGRGVGPGAGGAGDGT	692		
Db	524	prapsgrlqeraeqvsralqpaldsyfhpptgtpapgrgvpgagpgagdgdt	575		
RESULT 9					
AAV76143					
ID	AAV76143 standard; Protein; 309 AA.				
XX					
AC	AAV76143;				
XX					
DT	23-MAR-2000 (first entry)				
XX					
DE	Human secreted protein encoded by gene 20.				
XX					
KW	Human; secreted protein; cancer; tumour; developmental abnormality;				
KW	foetal deficiency; blood disorder; immune system disorder; inflammation;				
KW	autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;				
KW	schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;				
KW	atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;				
KW	digestive disorder; endocrine disorder; infection; AIDS; leukaemia;				
XX	therapy; chromosome 3.				
OS	Homo sapiens.				
XX					
PN	W09958660-A1.				
XX					
PD	18-NOV-1999.				
XX					
PF	06-MAY-1999; 99WO-US09847.				
XX					
PR	12-MAY-1998; 98US-0085093.				

PR 12-MAY-1998; 98US-0085094.
PR 12-MAY-1998; 98US-0085105.
PR 12-MAY-1998; 98US-0085180.
PR 18-MAY-1998; 98US-0085906.
PR 18-MAY-1998; 98US-0085920.
PR 18-MAY-1998; 98US-0085921.
PR 18-MAY-1998; 98US-0085922.
PR 18-MAY-1998; 98US-0085923.
PR 18-MAY-1998; 98US-0085924.
PR 18-MAY-1998; 98US-0085928.
PR 18-MAY-1998; 98US-0085925.
PR 18-MAY-1998; 98US-0085927.
XX
XX
XX
XX
PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
PI Olsen HS, Shi T, Young PE, Wei F, Brewer LA, Soppet DR;
PI Lafleur DM, Endress GA, Ebner R;
XX
XX
DR WPI; 2000-062296/05.
DR N-PSDB; AAZ65269.
XX
XX
PT New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -
PS
XX
XX
XX
XX
CC AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
CC AAY76124 to AAY76223 are the secreted proteins encoded by the 97 human
CC genes. The gene encoding this protein was found to be on chromosome 3.
CC The genes and their corresponding secreted polypeptides are
CC useful for preventing, treating or ameliorating medical conditions,
CC e.g. by protein or gene therapy. Also pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new genes. Specific
CC uses are described for each of the 97 genes, based on which tissues they
CC are most highly expressed in, and include developing products for the
CC diagnosis or treatment of cancer, tumours, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
CC disorders, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The sequences shown in AAY76224 to AAY76424 represent fragments of the
CC secreted proteins.
XX
XX
SQ Sequence 309 AA;

Query Match 38.5%; Score 1432; DB 21; Length 309;
Best Local Similarity 90.9%; Pred. No. 2.4e-122;
Matches 270; Conservative 0; Mismatches 11; Indels 16; Gaps 2;

QY	1	MPVPWFLLSLALGRSPVVLRLVGPQDATHCPSGLSRLWSDILCLPGDIVPAGPV	60
Db	1	mpvpwfllslalgrspvvlrlvvgpqdathcspglscrldwldilclpgdivpapgvp	60
QY	61	LAPTHLQTELVLRCQKEDCDLCLRVAVHLVHGHWEPEDEKFGGAADSGVEEPRNAS	120
Db	61	laphlqtelvlrcqketcldclrvxhvlavhghwepeedeekfggaadlgveeprnas	120
QY	121	LQAOVLVSFOAYPTARCVLLEVOVPALVQFGQSVGVVYDCFEALGSEVRINSYTQPR	180
Db	121	lqaovlvsfqaayptarcvllvqvpaaalvqfgqsvgvvydcfeaalgservlwsytqpr	180
QY	181	YEKELNHTQQLP-----ALPWLNVASADGDNVHLVNLNVEEQHFGLSLYWN	225
Db	181	yekelnhtqqipdcrglevwnslpwcwlpwlnvsadgdnvhlvlnvseeqnfqislywn	240
QY	226	QVQGPAPKPRWHKNTGPOIITLNTDLVPCICIQVPLEPDSVRTNICPFREDPRAH	282
XX			

Db 241 qvqgppkprwhkntgqitlnhtdlvpcicqvwplepdsrvrtapsgrtp-ah 296

RESULT 10
AAB88448
ID AAB88448 standard; Protein: 267 AA.
XX
AC AAB88448;
XX
DT 23-MAY-2001 (first entry)
XX
DE Human membrane or secretory protein clone PSEC0233.
XX
KW Human; secretory protein; membrane protein; vaccine; gene therapy;
KW rheumatoid arthritis; diabetes.
XX
OS Homo sapiens.
XX
PN EPI067182-A2.
XX
XX 10-JAN-2001.
XX
XX 07-JUL-2000; 2000EP-0114090.
XX
XX 08-JUL-1999; 99JP-01941179.
PR
XX 11-JAN-2000; 2000JP-0118775.
PR
XX 02-MAY-2000; 2000JP-0183766.
PR
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
XX
XX WPI; 2001-093989/11.
DR
DR N-PSDB; AAF93875.
XX
XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT gene therapy or as candidate target molecules in drug development -
PT
XX
XX Claim 1; SEQ ID 264; 609pp + CD ROM; English.
XX
XX This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC which encode human secretory or membrane proteins represented by
CC AAB88317 - AAB88419. Included in the invention are primers
CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
CC cDNA sequences of the invention. The invention also includes methods for
CC the production of antibodies directed against the proteins, and cDNA
CC sequences, which can be used in vaccines. The polynucleotide sequences
CC can be used in gene therapy. The polynucleotide sequences and the
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with inappropriate secretory
CC protein/membrane protein expression. The nucleic acids and complementary
CC sequences may also be used as DNA probes in diagnostic assays
CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
CC presence of similar nucleic acid sequences in samples. They may also be
CC used to study the expression and function of secretory proteins/membrane
CC polypeptides and their role in metabolism. The polypeptides may be used
CC as antigens in the production of antibodies against them and in assays to
CC identify modulators (agonists and antagonists) of expression and
CC activity. The antibodies and antagonists may also be used as therapeutic
CC agents to down regulate expression and activity. The antibodies may also
CC be used as diagnostic agents for detecting the presence of the
CC polypeptides in samples (e.g. by enzyme linked immunosorbent assay
CC (ELISA). Examples of diseases which may be treated include rheumatoid
CC arthritis and diabetes.
XX
SQ Sequence 267 AA;

Query Match 35.1%; Score 1305; DB 22; Length 267;
Best Local Similarity 89.2%; Pred. No. 8e-111;
Matches 249; Conservative 1; Mismatches 17; Indels 12; Gaps 2;

Oy 1 MPVPWFLLSLALGRSPVLSRLRVGPQDATHCSPGLSRLWDSIDLGLPGDIVPAGPV 60

Db 1 mpvpwfllslalgrspvlsrlrvgpqdashcsrglsrldlclpgdivpapgpv 60
QY 61 LAPTHLQTELVLRQCQKEDCDLCLRVAVHLAVHGHWEPEDEKFGGAADSGVEEPRNAS 120
Db 61 lapthlqtevlrvcqketcdclclrvavhlavhghweepedeekfggaadsgveeprnas 120
QY 121 LQAQVVLSTQAYPTARCVCVLELVQVPAALVQFGSGVSVYDCFEAALGSEVRWISYTPQR 180
Db 121 lqaqvvlstfqayptarcvcvllievqpaaalvqfgsgvsvvydcfeaalgsevrwisytcqr 180
QY 181 YEKELNHTQOLPALPWLNVSGADGNVHLVNLVSEEOHFGLSLYWNQYQGPKPWRHKNLT 240
Db 181 yekelnhtqqlpalpwlrvnsadgdnvhlvlnvseeqhfglslvwnvqgqgppkprwhknlv 240
QY 241 GPOITILNHTFDLVPCLCIQVMPLEPDSVRTNICPFREDP 279
Db 241 rpp-----psqvshsc-----rmpvpqrtqchiredp 267

RESULT 11
AAM25864
ID AAM25864 standard; Protein: 332 AA.
XX
AC AAM25864;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:1379.
XX
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnerary; antitumor; osteopathic; eczema;
KW dermatological; anti-allergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.
XX
OS Homo sapiens.
XX
XX WO200153455-A2.
PN
XX 26-JUL-2001.
XX
XX 22-DEC-2000; 2000WO-US35017.
PF
XX 23-DEC-1999; 99US-0471275.
PR
XX 21-JAN-2000; 2000US-0488725.
PR
XX 25-APR-2000; 2000US-0552317.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-457603/49.
XX
XX N-PSDB; AAM99805.
DR
XX
XX Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
PT
XX
XX Claim 20; Page 283; 1217pp; English.
PS
XX AAM99166 to AAM99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antinflammatory; antirheumatic;

CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; viricide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnerary;
CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy. Antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.

XX Sequence 332 AA;

Query Match 35.0%; Score 1300; DB 22; Length 332;
Best Local Similarity 86.4%; Pred. No. 3.1e-110;
Matches 248; Conservative 5; Mismatches 12; Indels 22; Gaps 3;

QY 1 MPVPFLLSLALGRSPVLSRLVGPQDATHCSGLSLWSDILCLPGDIVPAGPV 60
Db 2 MPVPFLLSLALGRSPVLSRLVGPQDATHCSGLSLWSDILCLPGDIVPAGPV 61
QY 61 LAPTHLOTLELRCORETDCDCLRVAVLAVHGHWEPEDEKFGCAADSGVEEPNAS 120
Db 62 lapthlqtelvlrcqketcdclrvavlavhghwepeedeekfgaadsveepnas 121
QY 121 LQAOVLSFQAYTARCVLLEQVPAALVQFGSGSVVYDCEAALGSGVRWISYQTPR 180
Db 122 lqaoavlsfqaytarcvllvqvpaaalvqfgsgsvvydcfeaalgsvevriwysyqtr 181
QY 181 YKELNHTQOLP-----ALPWLNVSDAGDNVHLVNLVNSEQHFGSLSYN 225
Db 182 yekelnhtqldqrglevnswpslcpwlnvsadgdnvhlvlnvseeqhfslsyn 241
QY 226 QVQGPKPWKNLNGP---QITLNTHTDLVPCLCITQVWPLEPDSVR 269
Db 242 qvqgpkprwhknlvrpppsqv---hshcrpclkcdavpyqrgslk 284

RESULT 12

AAE10920
ID AAE10920 standard; Protein; 204 AA.

XX AC AAE10920;

XX DT 18-DEC-2001 (first entry)

XX DE Human gene 12 encoded immune system-related protein HADEX14.

XX KW Human; immune system-related protein; allergy; rheumatoid arthritis;
KW cancer; Gaucher's disease; viral hepatitis; gene therapy; cytostatic;
KW diabetes mellitus; arrhythmia; wound healing; ischaemic lesion; AIDS;
KW Acquired Immune Deficiency Syndrome; viricide; hepatotropic; vasotropic;
KW autoimmune disorder; inflammation; cardiovascular disorder; hair loss;
KW wound healing; cell proliferation; skin aging; endocrine disorder;
KW food preservative.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Domain 50..55

XX FT /label= Immunogenic_epitope

XX FT Domain 151..168

XX FT /label= Immunogenic_epitope

XX

PN WO200166722-A1.

PD 13-SEP-2001.

XX 07-MAR-2001; 2001WO-US07260.

XX 08-MAR-2000; 2000US-187873P.

PR 11-AUG-2000; 2000US-224367P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Hilbert D, Kenny JJ, Moore PA, Choi GH, Soppet DR, Ebner R;
XX Gruber JR, Endress GA, Ruben SM;
XX WPI; 2001-589939/66.
XX N-PSDB; AAD18278.

XX Novel isolated immune system-related polypeptide useful for treating
XX rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease,
XX diabetes mellitus, arrhythmia, wound healing, ischemic lesions and
XX viral hepatitis

XX Claim 11; Page 312-313; 315pp; English.

XX The invention relates to human immune system-related protein and their
XX DNA. Human immune system related protein and DNA are useful for
XX preventing, treating or ameliorating a medical condition in a mammalian
XX subject, for diagnosing, preventing or treating immune system-associated
XX disorders, autoimmune disorders (rheumatoid arthritis), inflammatory
XX disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders
XX (allergies), infectious diseases (e.g., viral hepatitis), complement
XX activation disorders, immune complex diseases, neoplastic disorders
XX (cancer), hyperproliferative disorders (Gaucher's disease), disorders
XX associated with neovascularisation, diseases at the cellular level,
XX cardiovascular disorders (arrhythmias), wound healing and epithelial
XX cell proliferation, endocrine disorders (diabetes mellitus) and
XX neurological disorders (ischaemic lesions). Immune-system related protein
XX or DNA is useful for preventing hair loss, skin aging due to sunburn, to
XX maintain organs before transplantation, to treat weight disorders, to
XX modulate mammalian characteristics, to change a mammal's mental or
XX physical state, or as a food additive or preservative. Immune-system
XX related DNA is useful in gene therapy, for chromosome identification,
XX radiation hybrid mapping, long range restriction mapping and in forensic
XX biology. The present sequence represents a human immune-system related
XX protein of the invention.

XX Sequence 204 AA;

Query Match 29.0%; Score 1079; DB 22; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.5e-90;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 489 LLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRLSAGQPVAFWHAQRQTLEGGV 548
Db 1 llysaddsgferlvgalasalclqlplrvavdlwsrrelsagqpvafwfhqrqtllqegv 60

QY 549 VWLLFSPGAVALCSEWLQDVGSGGAGHPHDAFRASLSVLPDLQGRAPSVYGACFDR 608
Db 61 vwllfsgpavalcsewldqvgsggaghpghdafrafraslscvlpdlqgrapsgvygacfd 120

QY 609 LLHPDPAVPALFRTVPVFTLPSOLFALCQOPRAPSRLQERAEQVSRALQPALDSYF 668
Db 121 llhpdavpalfrtvpvftlpsqlpdfalqgqpraprsrlqeraeqvsralqpaldsyf 180

QY 669 HPGTTPAGRGVGPAGGAGDGT 692

Db 181 hpgtppagrgvgpaggagdgdt 204

RESULT 13

AA76048

ID AAY76048 standard; Protein; 617 AA.

XX AC AAY76048;
 XX AC
 XX DT 27-MAR-2000 (first entry)
 XX DE
 XX DE Murine skin cell protein, SEQ ID NO:303.
 XX KW Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KW secreted; transmembrane; inflammation; cancer; neurological disease;
 KW angiogenesis; tumour vascularisation; growth disorder;
 KW developmental disorder; skin wound; hair follicle disorder;
 KW anti-inflammatory; cytostatic; neuroprotective; vulnery.
 XX OS Mus sp.
 XX FN WO9955865-A1.
 XX PD 04-NOV-1999.
 XX PF 29-APR-1999; 99WO-NZ00051.
 XX PR 29-APR-1998; 98US-0069726.
 XX PR 09-NOV-1998; 98US-0188930.
 XX PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
 XX WPI; 2000-072177/06.
 XX DR N-PSDB; AAZ61753.
 XX PT Novel polynucleotides useful for the treatment of various conditions
 XX including wounds and cancer -
 XX PS Claim 4; Page 179-180; 235pp; English.
 XX CC The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,
 CC cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of
 CC cancer cells, to modulate angiogenesis and tumour vascularisation, to
 CC modulate skin inflammation, to modulate epithelial cell growth and to
 CC inhibit binding of HIV-1 to leukocytes. The invention may also be used
 CC to treat growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded
 CC by cDNA sequences derived from several mouse, rat or human skin cell
 CC types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and
 CC AAY76119 are proteins with an N-terminal signal sequence, indicating
 CC that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071,
 CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more
 CC putative transmembrane domains.
 XX Sequence 617 AA;
 XX SQ

Query Match 8.8%; Score 326.5; DB 21; Length 617;
 Best Local Similarity 26.7%; Pred. No. 8.7e-21;
 Matches 132; Conservative 56; Mismatches 184; Indels 123; Gaps 19;

QY 237 KNLGTGQIITLHNTDLVPCICIQWVPLEPDSVRTNICPFREDPRAH-QNLWQAAR----- 290
 DB 183 kivsghvtldpyefilpcmcieasyldqetvrrkkcpqgwpqagsgdfwqsfirtdys 242
 QY 291 -----LRLTLQSWLLDAPCSLPAAALCWAPGPGPCQPLVPLSWENTVTVKVLFFPL 345
 DB 243 qhngmvmaltlr-----cpkleaslcwrgdpltpcetl-----pnataqesegwyi 289
 QY 346 LAG---HPNLCVQVNSSEKLOEQECLWADSLGPLKDDVLLLETRGPDNRSKALEPSGC 402
 DB 290 lenvdlhpgqlcfkf-sfensshvec--phqsgslpwtvtsmdtgaqg-----ltlhfsr 341

QY 403 TSLPSKASTRAARLG-----EYLLQDLQSGCLOLWDDDLGAL 440
 DB 342 tyatfsaawsdpglpgdtpmpvysisqtgsvptldliipfrqencilivrsdvhfa 401
 QY 441 W---ACPMKYIHKRWALVWLACLLFAAALSILILLLKKDHAKAAAGRAALLLYSADDSG 497
 DB 402 wkhvlpdd-----apyptqlllr-----slsggrtrpvlillhaadsea 440
 QY 498 FERLVGALASALCOLPRLVA-----VDLWSRRELSAQGPVAFWFAHORRQTLOEGGVV 549
 DB 441 qrrivgalael-----lrtalggdrdvldwegthvarigpplwlaarervaregtv 495
 QY 550 VLLFSPGVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLOGRAPGSGVYGACFRL 609
 DB 496 lllwn-----cagptacsgdpqaaslrtil-----caaprplllayfsrl 536
 QY 610 LHPDAVPALFRTPVETLPSOLPDFLGALO-QPRAPRSG-----RLQRAEOVS-RA 659
 DB 537 cakgdiprprlralprylrrldprllraldaqpattlasswshlgakrcldknrlqechlle 596
 QY 660 LQPALDSYFHPPGTP 674
 DB 597 leaakddyggstnsp 611

RESULT 14

AAB55987
 ID AAB55987 standard; Protein; 617 AA.
 XX AC AAB55987;
 XX DT 08-MAR-2001 (first entry)
 XX DE Skin cell protein, SEQ ID NO: 303.
 XX KW Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;
 KW neurotropic; neuroprotective; vulnery; immunomodulatory; vaccine;
 KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
 KW inflammation; neurological disease.
 OS Mus sp.
 XX WO200069884-A2.
 XX PD 23-NOV-2000.
 XX PF 15-MAY-2000; 2000WO-NZ00075.
 XX PR 14-MAY-1999; 99US-0312283.
 XX PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX PI Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD, Murison JG;
 XX WPI; 2001-007495/01.
 XX DR N-PSDB; AAC99686.
 XX PT New isolated polynucleotide used in the identification of genetic
 PT disorders and encoding polypeptides used for treating inflammatory
 XX disease, cancer and neurological diseases -
 XX PS Claim 4; Page 244-245; 352pp; English.
 XX CC The present sequence is a polypeptide which is expressed in
 CC mammalian skin cells. The polypeptide is useful for stimulating
 CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
 CC modulating angiogenesis, inhibiting angiogenesis and vascularisation of
 CC tumours, modulating skin inflammation, stimulating the growth of
 CC epithelial cells, inhibiting the binding of human immunodeficiency virus
 CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
 CC neurological diseases. The polynucleotide can be used as a marker, in
 CC the identification of genetic disorders, and for the design of

CC degenerative cartilaginous disorder comprises administering a PRO1031 or
 CC PRO1122 polypeptide agonist, or antagonist to the mammal. Numerous
 CC examples of the diseases and disorders are given in the specification.

XX
 SQ Sequence 667 AA;

```

Query Match      8.3%; Score 309.5; DB 22; Length 667;
Best Local Similarity 29.0%; Pred. No. 3.5e-19;
Matches 138; Conservative 54; Mismatches 181; Indels 103; Gaps 23;

Qy 237 KNLGPGQITLNNHTDLPVPCICIQVWPLEPDSVTNCPREDPRAH-QNLWQAAAR----- 290
Db 223 kivsghgtvelpfeilpcleasylqetvrrkkcpqswpeaygsdfwksvhtfids 282
Qy 291 -----LRLTLQSWLLDAPCSLPAEALCWAPGPGCPQPLVPPLSWENVTVDK-----V 340
Db 283 qhtqmvmaaltf-----cpkleaalqrhdwhlckdl-----pnataresdgwyv 329
Qy 341 LEFPLKGNHLCVQV---NSSEKIQEQELWADSLGPKDDVLLLETRGPQ-----DNR 392
Db 330 lekvdI--hpqlcfkfsfgnssh-----vec--phqtgsitwnvsmtdgaqllilhfssr 381
Qy 393 SLCALEPSGCTSLP-----SKASTRAARLGEYLLQDLQSGQCLQLWDDDLG 438
Db 382 mhatf--saawslpglqgdtlvppvytvsgargsspsvldliipfipgcvlvwrsdvq 439
Qy 439 ALW---ACPMDKYIHKRWALVWLACILFAAALSILLLK-KDHAKAAARGRAALLLYSAD 494
Db 440 fawkhlilcpdvsvyrh--lgllilallliltlgvialtcrrpqsppgparpvlllhaad 497
Qy 495 DSGFERLVGALASALCQLPLRVA-----VDLWSRRELSAQGPVAFHQAQRRTIQEG 546
Db 498 seaqrllvgalael-----lraalggrdvivdlwegrhvarvgpplwlaartvareq 552
Qy 547 GVVVLLFSPGVALCSEWLQDGVSGGAHGHDAFRASLSCLVLPDFLQGRAGS-VYGAC 605
Db 553 glvlllwsqadlr-----pvsqp-----dpraaplall-----haaprlillay 593
Qy 606 FDRLLHPDAVPALFRTVPVFTLPSQLPDLFGALQ-QPRAPRS--GRLQERAEQVSR 658
Db 594 fsrlcakgdiippalpryllrdlprllraldarpfaeatswgrlqarqrqsr 649
  
```

Search completed: May 19, 2002, 13:26:40
 Job time: 7621 sec